

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: May 13, 2006, 09:34:18 ; Search time 189.5 Seconds  
(without alignments)  
4953.591 Million cell updates/sec

Title: US-10-701-844-1  
Perfect score: 7883  
Sequence: 1 999caaaactctccccccg.....gcactcaacgaagaattc 4435

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2165443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DRV=xlp  
-Q=/abs/ABSSWEB spool/US10701844/runat 12052006 165420 26277/app query.fasta\_1  
-DB=UniProt -QEMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USER=US10701844 @CGN 1.1 466 @runat 12052006 165420 26277 -NCPU=6 -ICPU=3  
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLAG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_tr embl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5131.5	65.1	1013	1 PMPG CHLTR	O84879 chlamydia t
2	3845	48.8	987	1 PMPG CHLMU	O9p145 chlamydia m
3	1860.5	23.6	1011	2 Q823X5 CHLCV	O823x5 chlamydothi
4	1777	22.5	1024	2 Q5L6J3 CHLAB	Q5L6j3 chlamydothi
5	1589	20.2	1016	1 PMPH CHLTPN	O84880 chlamydia t
6	1566	19.9	981	2 Q84FV7 CHLTR	O84fu7 chlamydia t
7	1551.5	19.7	980	2 Q84FV6 CHLTR	O84fu6 chlamydia t
8	1551.5	19.7	980	2 Q84FV8 CHLTR	O84fu8 chlamydia t
9	1459	18.5	989	2 Q83U76 CHLTR	O83u76 chlamydia t
10	1459	18.5	989	2 Q84FV9 CHLTR	O84fu9 chlamydia t
11	1452	18.4	991	2 Q83UW2 CHLTR	O83uw2 chlamydia t
12	1452	18.4	991	2 Q84FV0 CHLTR	O84fv0 chlamydia t
13	1452	18.4	991	2 Q84FV1 CHLTR	O84fv1 chlamydia t
14	1404.5	17.8	993	2 Q83TJ6 CHLTR	O83tj6 chlamydia t
15	1404.5	17.8	993	2 Q84FV2 CHLTR	O84fv2 chlamydia t
16	1365.5	17.3	973	1 PMP13 CHLPN	Q92896 chlamydia p

17	1174	14.9	980	1 PMPH CHLMU	Q9p144 chlamydia m
18	1158	14.7	942	2 Q823X1 CHLCV	Q823x1 chlamydothi
19	1122	14.2	926	2 P71135 CHLAB	P71135 chlamydothi
20	1122	14.2	926	2 Q5L6J3 CHLAB	Q5L6j3 chlamydothi
21	1101	14.0	928	1 PMP10 CHLPN	Q92b65 chlamydia p
22	1099	13.9	926	2 Q823W9 CHLCV	O823w9 chlamydothi
23	1090	13.8	928	1 PMP11 CHLPN	O86164 chlamydia p
24	1079	13.7	930	1 PMP8 CHLPN	Q92393 chlamydia p
25	1077	13.7	1378	2 Q5L6J2 CHLAB	Q5L6j2 chlamydothi
26	1040	13.2	936	1 PMP7 CHLPN	Q92898 chlamydia p
27	1021	13.0	928	1 PMP9 CHLPN	Q92398 chlamydia p
28	1015	12.9	1276	1 PMP6 CHLPN	Q92899 chlamydia p
29	1008	12.8	866	2 Q823X0 CHLCV	Q823x0 chlamydothi
30	979.5	12.4	772	2 Q9RB71 CHLPN	Q9rb71 chlamydia p
31	973	12.3	922	1 PMP1 CHLPN	Q92895 chlamydia p
32	971.5	12.3	841	2 Q822Q5 CHLCV	Q822q5 chlamydothi
33	950	12.1	847	2 P71132 CHLAB	P71132 chlamydothi
34	949.5	12.0	839	2 P77792 CHLAB	P77792 chlamydothi
35	939.5	11.9	841	1 PMP2 CHLPN	Q923a1 chlamydia p
36	897	11.4	846	2 P71133 CHLAB	P71133 chlamydothi
37	803	10.2	843	2 Q823X2 CHLCV	Q823x2 chlamydothi
38	789	10.0	867	1 PMP1 CHLMU	Q9p141 chlamydia m
39	787.5	10.0	602	2 Q8VU49 CHLPS	Q8vu49 chlamydia p
40	785.5	10.0	843	2 Q823X3 CHLCV	Q823x3 chlamydothi
41	785	10.0	849	2 Q5L6J5 CHLAB	Q5L6j5 chlamydothi
42	784.5	10.0	878	2 Q4VWS6 CHLTR	Q4vws6 chlamydia t
43	781.5	9.9	878	2 Q4VWR2 CHLTR	Q4vwr2 chlamydia t
44	779.5	9.9	878	2 Q4VWS3 CHLTR	Q4vws3 chlamydia t
45	779.5	9.9	878	2 Q4VWS8 CHLTR	Q4vws8 chlamydia t

## ALIGNMENTS

RESULT 1  
PMPG\_CHLTR  
ID PMPG CHLTR STANDARD; PRT; 1013 AA.  
AC O84879;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G).  
DE Name=pmpG; OrderedLocusNames=CT871;  
GN Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
ON NCBI\_taxid=813;  
RX [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=D/UV-3/Cx;  
RX MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;  
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;  
RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";  
RT Science 282:754-759(1998).  
RL -|- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (potential).  
CC -|- SIMILARITY: Belongs to the PMP outer membrane protein family.  
CC -|- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC ENBL; A8001360; AAC68469.1; -; Genomic\_DNA.  
DR PIR; G71460; G71460.  
DR PHCI-2DPAGE; O84879; -  
DR InterPro; IPR005546; Auto\_transpbtbeta.  
DR InterPro; IPR006315; Autotransporter.  
DR InterPro; IPR011427; ChlamPMP\_M.

DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 9.  
 DR Pfam; PF07548; ChlamPMP\_M; 1.  
 DR TIGRFAMs; TIGR01414; autotrans\_bar1; 1.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 5.  
 KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.  
 FT SIGNAL 1 27 Potential.  
 FT CHAIN 28 1013 Probable outer membrane protein pmpG.  
 SQ SEQUENCE 1013 AA; 107367 MW; F0927743C0A651DD CRC64;

## Alignment Scores:

Pred. No.: 0 Length: 1013  
 Score: 5131.50 Matches: 983  
 Percent Similarity: 98.8% Conservative: 18  
 Best Local Similarity: 97.0% Mismatches: 11  
 Query Match: 65.1% Indels: 1  
 DB: 1 Gaps: 1

US-10-701-844-1 (1-4435) x PMPG\_CHL1TR (1-1013)

QY 382 ATGCAAGCTCTTCCATAAGTCTCTTCTTCAATGATTCTAGCTTATTCTTGCTGCTCT 441  
 DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20  
 QY 442 TTAATGGGGGGGATGCGAGCAAAATCATGTTCTCTCAAGGAATTATCATGGGGAG 501  
 DB 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
 QY 502 ACGTTAACTGATCATTTCCCTATCTCTTATAGGATCCGAGTGGGACTGTTTTT 561  
 DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
 QY 562 TCTGAGGAGAGTTAACTTAAATACTTGACAATCTTATGACAGCTTGTGCTTTAAGT 621  
 DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80  
 QY 622 TGTTTGGAACTTATTAGGGAGTTTACTGTGTTTGGAGAGGACACTCGTTGACTTTC 681  
 DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 QY 682 GAGACATACGACTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741  
 DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
 QY 742 TTTACTATTGAGGGTTTAAAGAAATTATCTTTTCCAAATTCGAATTCATTACTTGCCTGA 801  
 DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
 QY 802 CTGCTGCTGCACCACTAATAAGGTAGCCAGACTCCGAGCAACAATCTACACCGTCT 861  
 DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160  
 QY 862 AATGCTACTATTATTCTTAAACAGACTCTTTTGTACTCAATATGAGAAGTCTCATTC 921  
 DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
 QY 922 TATAGTAATTAGTCTCTGGAGATGGGAGCTATAGATGCTTAAAGCTTAAACGGTTCAA 981  
 DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
 QY 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
 DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
 QY 1042 CAACTAGTACCAAGTTCTCTGCTATGCTTAACGAGGCTCTATTGCTTGTAGCGAAT 1101  
 DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
 QY 1102 GTTGAGGAGTAAGAGGGGAGGATGCTGCTGTTCAGGATGGGAGGAGGTGCA 1161  
 DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260

QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACTCGCGTAGAGTTT 1221  
 DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
 QY 1222 GATCGGAACGTAGCCCGAGTAGGAGGAGGATTACTCTCTACGGGAACGTTGCTTCTCTG 1281  
 DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
 QY 1282 AATAATGGAAAAACCTTTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAG 1341  
 DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320  
 QY 1342 CAACCAACAGTGCAGCAGGCTTCTAATACGATTAATTAATACGAGATGAGGAGCTATC 1401  
 DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
 QY 1402 TTCTGTAAGAATGTCGCGCAA---GCAGGATCCAAATACTCTCGATCAGTTTCTTCTGAT 1458  
 DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
 QY 1459 GGAGAGGAGTAGTTTCTTTTAGTACAAATGTAGTCTCTGGGAAAGGGGAGCTATTAT 1518  
 DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380  
 QY 1519 GCCAAAAAGCTCTCGGTTGCTAAGTCTGCGCTGTACAAATTTTAAAGAAATATCGCTAAT 1578  
 DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
 QY 1579 GATGCTGAGCGATTTATTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
 DB 401 AspGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420  
 QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACACCCAAAGAGAAATGCTGCCGATGTTAAT 1698  
 DB 421 AspileIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaLeaAspValAsn 440  
 QY 1699 GGCCTAACTGTGCTCTCAAGCCATTTCCGATGGGATCGGGAGGAGGAAAAATAACGACATTA 1758  
 DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
 QY 1759 AGCCTAAAGCAGGCGATCAGATTCCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818  
 DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
 QY 1819 AACCAGCAGCGCAGTCTTCCAACTTCTAAATAATAACGATGCTGAAGGATACACAGGG 1878  
 DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyGlyTyrThrGly 500  
 QY 1879 GATATTGTTTGTCTAATGAGCAGTACTTTGTACCAAAATGTTACGATAGACAGGA 1938  
 DB 501 AspileValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
 QY 1939 AGGATTGTTCTCTGTAAGGCAAAATTAATCAGTGAATTTCTAAGTCAGACAGGTGGG 1998  
 DB 521 ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
 QY 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCAACCAACAG 2058  
 DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
 QY 2059 CCTCTGCGCTAATCAGTTGATCACGCTTTCGAATCTGCAATTTGCTCTTCTTCTTCTTG 2118  
 DB 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
 QY 2119 TTAGCAACAATGAGTTAGCAATCTCTTACCAATCTCCAGCGCAAGATTTCTCATCCT 2178  
 DB 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
 QY 2179 GCAGTCAATTGTCACAACTGCTGTTTCTGTTTACAATTAGTGGGCGCTTATCTTTTTCAG 2238  
 DB 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
 QY 2239 GATTGGATGATCAGCTTATGATAGGTATGATTGGGTAGTCTTCTAATCAAAAAATCAAT 2298



Db	621	AspLeuAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnIysIleAsp	640		981	SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer	1000
QY	2299	GTCTCTGAAATTACAGTTAGGACTAAGCCCCAGCTTAATGCCCCATCAGATTGACTCTA	2358		3379	CGAGCTATGTTTTCAGTGACGAGGAAGTAGAGTCCGGTTC	3417
Db	641	ValLeuIysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660		1001	ArgGlyTyrGlyLeuSerAlaGlySerIysValArgPhe	1013
QY	2359	GGGAATGAGATCGCTTAAGTATGGCTATCAAGNAGCTGGAAGCTGCGGGATCCTAAT	2418		RESULT 2		
Db	661	GlyAsnGluMetProIysTyrGlyTyrGlnGlySerTrpIysLeuAlaTrpAspProAsn	680		PMMPG_CHLMU	STANDARD;	PRT; 987 AA.
QY	2419	ACAGCAAAATAGTCTTATCTCTGAAAGCTACATGAGCTAAGAACTGGGTATATCTCT	2478		ID	PMMPG_CHLMU	987 AA.
Db	681	ThrAlaAsnGlyProTyrThrLeuIysAlaThrTrpThrIysThrGlyTyrAsnPro	700		AC	Q9PL45;	
QY	2479	GGGCTCTGAGCGAGTCTCTTCTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATA	2538		DT	16-OCT-2001	(Rel. 40, Created)
Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle	720		DT	16-OCT-2001	(Rel. 40, Last sequence update)
QY	2539	CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGGGCTCTTATTGTGAGGATTA	2598		DT	13-SEP-2005	(Rel. 48, Last annotation update)
Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740		DE	Probable outer membrane protein pmg precursor (Polymorphic membrane protein G).	
QY	2599	TGGGTTTCTGGAGTTTCGAATTTCTCTATCATGACCGGATGCTTTAGTTCAGGATAT	2658		GN	Name=pmgG; OrderedLocusNames=TC0263;	
Db	741	TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760		OS	Chlamydia muridarum.	
QY	2659	CGGTATATTAGTGGGGTTATTCTCTTAGGAGCAAACTCCCTACTTTGGATCATCGATTTT	2718		OC	Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydia.	
Db	761	ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780		OX	NCBI_TaxID=83560;	
QY	2719	GGTCTAGCATTTACCGAGTATTGTGTAGATCTAAGATTTATGCTGTCTGCTCCCAT	2778		RN	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	
Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerIysAspTyrValValCysArgSerAsn	800		RP	STRAIN=MoPr / Ni99;	
QY	2779	CATCATCTGTGATAGTCCGTTTCTATCTATCTACCAACAGCTTTATGTGATCTCTAT	2838		RC	MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;	
Db	801	HisHisAlaCysIleGlySerValTyrLeuSerThrIysGlnAlaLeuGlySerTyr	820		RX	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,	
QY	2839	TTGTTCGGAGATCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATCAAAACC	2898		RA	White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,	
Db	821	LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr	840		RA	Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,	
QY	2899	TCATATACATTTGCAGAGGAGCGATGTTCTGTTGGGATTAATACCTCTCGGCTGGAGAG	2958		RA	Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,	
Db	841	SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu	860		RA	McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;	
QY	2959	ATTGGACGGGATTACCGATTGTGATTCTCCATCTAAGCTCTATTGTAATGATGTCGCT	3018		RT	"Genome sequences of Chlamydia trachomatis MoPr and Chlamydia pneumoniae AR39.";	
Db	861	IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg	880		RL	Nucleic Acids Res. 28:1397-1406(2000).	
QY	3019	CTTTCTGCGAAGCTGATTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT	3078		CC	-1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).	
Db	881	ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp	900		CC	-1- SIMILARITY: Belongs to the PMP outer membrane protein family.	
QY	3079	CAAGCTCGGCGATTCAAGAGCGGCATCTCTCTAAATCATCAGTCTCTGTTGGAGTCAAG	3138		CC	-1- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.	
Db	901	GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValIys	920		CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.	
QY	3139	TTTGATCATGTTCTTAGTACATCTCTAATAATATAGCTTTATGGCGCTTATATCTGT	3198		CC	EMBL; A2002293; AAF39132.1; -; Genomic DNA.	
Db	921	PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaIleCys	940		DR	PIR; H81722; H81722.	
QY	3199	GATGTTTCCACCATCTCTGGTACTCAGACAAAGCTCTCTATCCCATCAAGACATGG	3258		DR	TIGR; TC0263;	
Db	941	AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp	960		DR	InterPro; IPR005546; Auto_transptbeta.	
QY	3259	ACAACAGATGCTTTTCATTTAGCAAGACATCGAGTTGGTTAGAGATCTATATGCT	3318		DR	InterPro; IPR006315; Autotransporter.	
Db	961	ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla	980		DR	InterPro; IPR011427; ChlamPMP_M.	
QY	3319	TCTCTAACAGTATATAGAGTATATGCGCATGGAAGATATGATATCGAGATCTCTCT	3378		DR	InterPro; IPR003368; Chlamydia PMP.	
					DR	Pfam; PF03797; Autotransporter; 1.	
					DR	Pfam; PF02415; Chlam_PMP; 9.	
					DR	Pfam; PF07548; ChlamPMP_M; 1.	
					DR	TIGRFAMS; TIGR01414; autotrans_bar1; 1.	
					DR	TIGRFAMS; TIGR01376; POMP repeat; 5.	
					KW	Complete proteome; Membrane; Multigene family; Outer membrane; Signal.	
					FT	SIGNAL	1 25 Potential.
					FT	CHAIN	26 987 Probable outer membrane protein pmgG.
					SQ	SEQUENCE	987 AA; 104867 MW; 32079BD6BEB2DA42 CRC64;
				Alignment Scores:			
				Pred. No.:	1.28e-246	Length:	987
				Score:	3846.00	Matches:	739
				Percent Similarity:	82.9%	Conservative:	106
				Best Local Similarity:	72.5%	Mismatches:	136
				Query Match:	48.8%	Indels:	38
				DB:	1	Gaps:	13
				US-10-701-844-1 (1-4435) x PMMPG_CHLMU (1-987)			
QY	379	GTGATGCAACGCTTTTCCATAAGTCTTTCTTTCAATGATCTAGCTTATCTTCTGCTGC	438				
Db	1	MetMetGlnThrProPheHisPhePheLeuLeuAlaMetLeuSerTyr	17				



Db 709 AlaTyrCysArgGlyIleTrpIleSerGlyIleSerAsnPhePheTyrHisAspGlnAsp 728  
 QY 2641 GCTTTAGGTCAGGATATCGGTATATATAGTGGGGGTTATTCCTTAGGACCAATCCTCAC 2700  
 Db 729 AlaLeuGlyGlnGlyTyrArgHisIleSerGlyGlyTyrSerIleGlyAlaAsnSerTyr 748  
 QY 2701 TTTGGATCATCGATGTTGGTCTAGCATTTACCAGATATTGGTAGATCTAAAGATTAT 2760  
 Db 749 PheGlySerSerMetPheGlyLeuAlaPheThrGluThrPheGlyArgSerIleAspTyr 768  
 QY 2761 GTATGTGTGCTTCCAAATCATATGCTTCATAGATCCGTTTATCTATCTACCCACAA 2820  
 Db 769 ValValCysArgSerAsnAspHisThrCysValGlySerValTyrLeuSerThrArgGln 788  
 QY 2821 GCTTTATGTGGATCTTATTTGCGAGATGCGTTTATCCGTGCTAGCTACGGTTTGG 2880  
 Db 789 AlaLeuCysGlySerCysLeuPheGlyAspAlaPheValArgAlaSerTyrGlyPheGly 808  
 QY 2881 AATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGATGCTGTTGGGATAAT 2940  
 Db 809 AsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAsnValArgTyrAspAsn 828  
 QY 2941 AACTGTCGCTGCGAGATGCGAGGATGCGGATACCGATTGTGATTACTCCATCTAGCTC 3000  
 Db 829 AsnCysValValGlyGluValGlyAlaGlyLeuProIleMetLeuAlaAsnSerLysLeu 848  
 QY 3001 TATTTGAATGAGTGGTCTCTTTCGTCAGAGTGGATTTTCTTATCCGATCATGAATCT 3060  
 Db 849 TyrLeuAsnGluLeuArgProPheValGlnAlaGluPheAlaTyrAlaGluHisGluSer 868  
 QY 3061 TTTCAGAGAGGCGCATGAGCTCGGCGATTCAGAGCGGACATCTCTTAATCTATCA 3120  
 Db 869 PheThrGluArgGlyAspGlnAlaArgGluPheLysSerGlyHisLeuMetAsnLeuSer 888  
 QY 3121 GTTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTT 3180  
 Db 889 IleProValGlyValLysPheAspArgCysSerSerLysHisProAsnLysTyrSerPhe 908  
 QY 3181 ATGGCGGCTTATATCTGTATGCTTATCGCACCATCTCTGGTACTGAGACACGCTCCTA 3240  
 Db 909 MetGlyAlaTyrIleCysAspAlaTyrArgSerIleSerGlyThrGluThrThrLeuLeu 928  
 QY 3241 TCCCATCAAGACATGAGACACAGATGCTTCTTATTTACAGACATGAGTGTGGTT 3300  
 Db 929 SerHisLysGluThrTrpThrAspAlaPheHisLeuAlaArgHisGlyValMetVal 948  
 QY 3301 AGAGGATCTATGATGCTCTCTAACAAGTAATATAGATATATGATGCGCATGGAAGATAT 3360  
 Db 949 ArgGlySerMetTyrAlaSerLeuThrGlyAsnIleGluValTyrGlyHisGlyLysTyr 968  
 QY 3361 GAGTATCGAGATGCTCTCCAGGCTATGTTGATGTCAGAGCATGAGTCCGCTTC 3417  
 Db 969 GluTyrArgAspAlaSerArgGlyTyrGlyLeuSerIleGlySerLysIleArgPhe 987

RESULT 3  
 Q823X5\_CHLCV  
 ID Q823X5\_CHLCV PRELIMINARY; PRT; 1011 AA.  
 AC Q823X5\_2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DT Polymorphic outer membrane protein G family  
 DE protein/autotransporter.  
 GN OrderedLocusNames=CCA00278;  
 OS Chlamydia caviae.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.  
 OC NCBI\_TaxID=83557;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=GPIC;  
 RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;  
 RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,  
 Heidelberg J.F.; Holtzapfel E.K., Khouri H.M., Federova N.B.,

RA Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,  
 White O., Salzberg S.L., Haia R.-C., McClarty G., Rank R.G.,  
 Bavell P.M., Fraser C.M.;  
 RT "Genome sequence of Chlamydia psittaci GPIC":  
 RT examining the role of niche-specific genes in the evolution of the  
 RT Chlamydiaceae.";  
 RL Nucleic Acids Res. 31:2134-2147 (2003).  
 DR EMBL; AB016995; AAP05029.1; -; Genomic\_DNA.  
 DR TIGR; CCA00278; -;  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR InterPro; IPR006315; Auto transporter.  
 DR InterPro; IPR005546; Auto transportbeta.  
 DR InterPro; IPR011427; ChlamPMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF07548; ChlamPMP\_M; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 7.  
 DR TIGRFAMs; TIGR01414; autotrans\_bar1; 1.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 6.  
 KW Complete proteome.  
 SQ SEQUENCE 1011 AA; 107952 MW; EBD9A9F1C46E8E6 CRC64;

Alignment Scores:  
 Pred. No.: 9,73e-115 Length: 1011  
 Score: 1860.50 Matches: 426  
 Percent Similarity: 55.2% Conservative: 168  
 Best Local Similarity: 39.6% Mismatches: 353  
 Query Match: 23.6% Indels: 129  
 DB: 2 Gaps: 27

US-10-701-844-1 (1-4435) x Q823X5\_CHLCV (1-1011)  
 QY 382 ATGCAACGCTCTTCCATTAAGTCTTCTTCTTCA-----ATGATTTCTAGCTTATTTCTTGC 435  
 Db 1 MetLysAlaSerLeuArgLysPheLeuIleSerThrThrLeuThrLeuProTyrSerPhe 20  
 QY 436 TGCTCTTTAAATGGGGGGATATGACGACGAAATCATGCTTCTCCTCAAGGAATTTACGAT 495  
 Db 21 -----GlnAlaPheSerLeuGluValValProAsnGlyThrThrAsp 35  
 QY 496 GGGGAGACGTTAACTGATCTATCTTCCCTATCTGTTTATAGGAGATCCGAGTGGGACTACT 555  
 Db 36 Gly--AsnLeuArgGluThrPheProTyrThrIleThrSerAsnProGluGlyThrThr 54  
 QY 556 GTTTTTCTGCAGGAGATTAACATTTAAAAATCTTGACATTTCTATGTCAGCTTTTCCCT 615  
 Db 55 AlaIleLeuSerGlyAsnLeuAsnLeuLeuAsnLeuAspAsnSerMetValAlaThrPro 74  
 QY 616 TTAAGTTGTTGGGAACTTATTAGGAGTGTCTTCTGTTTATAGGAGGAGGAGACTCGTTG 675  
 Db 75 SerSerCysPheAsnSerAlaGlySerMetThrIleValGlyArgAsnHisAsnLeu 94  
 QY 676 ACTTTCGAGAACATACGACTTCTACAAATGGGGAGCTCTTAAGTAAT----- 723  
 Db 95 ThrPheThrAsnLeuArgThrSerAlaAsnGlyAlaAlaLeuSerSerIleProThrThr 114  
 QY 724 AGCGTGTGATGAGCTGTTTACTTATGAGGTTTAAAGAAATATCTTTTCCATTTCC 783  
 Db 115 ThrProGluSerPheProTyrThrIleLysGlyValAsnThrLeuSerPheSerAsnCys 134  
 QY 784 AATTCATTACTTCCGCTGCTGCTGCAACGACCTAATAAGGTAGCCAGACTCCGACG 843  
 Db 135 -----LeuAlaLeuMetAlaArgThrThrAlaProAsnThrThrThrProVal 151  
 QY 844 ACAACATCTACACGCTCTAATGCTACTATTATTCTTAAACACAGATCTTTTGTACTCAAT 903  
 Db 152 -----AsnProAsnGlyGlyAlaPheTyrSerLysAlaProValPheLeuGluAsn 168  
 QY 904 AATGAGAGTCTCTATTCTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCT 963  
 Db 169 IleGlnAsnValLeuPheLysAsnAsnArgAlaAlaAspSerGlyGlyGlyLeuTyrVal 188  
 QY 964 AAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTCTCCAGGAATAACTGCTCAA 1023

```

Db      189  GluThrAlaGlyIleSerAsnIleLysLysSerMetGlnPheLeuSerAsnValGly--- 207
      : : : : :
Qy      1024 GCTGATGGGAGCTGTCTCAAGTAGTACACAGTTTCTCTGCTATGCTTAACAGAGCTCCT 1083
      : : : : :
Db      208  AlaAsnGlyGlyAla-----IleAsnAlaSerLysSerLeuAspValThrGlnCysPro 225
      : : : : :
Qy      1084 ---ATTGCTTTGTAGCGAATGTTCCAGAGTAGTAAGAGGGGAGGAGTGTCTGCTGTTCCAG 1140
      : : : : :
Db      226  SerIleLeuPheArgSerAsnSerAlaGluLysLeuGlyGlyAlaIleGlnAlaValAsp 245
      : : : : :
Qy      1141 GATGGGACAGGGAGGTGCTCATCTACTTCAACAGAGAGATCCAGTAGTAGTAAAGTTTTC 1200
      : : : : :
Db      246  -----ProAlaThrThrAsnGlnValAsnThrAlaValArgPheSer 259
      : : : : :
Qy      1201 AGAATACTCGGTAGATGTTGATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCC 1260
      : : : : :
Db      260  GluAsnGlySerValGlnPheAspAlaAsnAsnAlaLysSerGlyGlyAlaIleTySer 279
      : : : : :
Qy      1261 TACGGGAAGTGTCTTCTCAATATGGAAGAACTGTTTCTCACAATGTTGCTTCT 1320
      : : : : :
Db      280  LysGlyAsnValAspPheSerAsnAsnAlaGlnLeuLeuIleGlnAsnAsnSerAlaSer 299
      : : : : :
Qy      1321 CCTGTTTACATTGCTCTAAGCAACCAAGTGACAGGCTTCTAATACAGTAATAAT 1380
      : : : : :
Db      300  Pro-----GluValAlaAsnThrAsnGluVal 308
      : : : : :
Qy      1381 TACGAGATGAGGAGCTATCTTCGT----- 1407
      : : : : :
Db      309  LeuGlyGlnGlyGlyAlaIlePheCysValGlnGlnThrProThrGlnProProPro 328
      : : : : :
Qy      1408 -----AAGATGGTGGCAACAGCAGATCCATAACTCTGGATCAGTTTCTCTT 1455
      : : : : :
Db      329  ProProProThrThrAsnProValPheSerGly-----LeuThrIle 342
      : : : : :
Qy      1456 GATGGAGAGGAGTAGTCTTTTAGTAGCAATGTAGTGTGGGAAAGGGAGCTATT 1515
      : : : : :
Db      343  ThrAsnGlnLysAspIleLeuPheAlaAsnAsnPheAlaAlaThrAlaGlyGlyAlaIle 362
      : : : : :
Qy      1516 TATGCCAAAAGCTCTCGTGTCTAAGTGTGGCCCTGTACAAATTTTAAAGAAATATCGCT 1575
      : : : : :
Db      363  TyrGlyGluLysValSerIleThrSerSerGlyLysThrMetPheThrAsnAsnIleAla 382
      : : : : :
Qy      1576 AATGATGGTGGAGATTTATTTAGGAATCTCGAGAGCTCAGTTTATCTGCTGATAT 1635
      : : : : :
Db      383  LysAspGlyGlyAlaIleTyIleProGluAsnGlyGluLeuThrLeuSerAlaAspTy 402
      : : : : :
Qy      1636 GGAGATATATTTCGATGGGAATCTTAAAGAAACAGCAAGAGATGCTGCGCATGTT 1695
      : : : : :
Db      403  GlyAspMetIlePheTyThrGluAsnLeuLysLys----- 413
      : : : : :
Qy      1696 AATGGCGTAACTGTGCTCTCAAGCCATTTTCGATGGATCGGGAGGAAATAACAGCA 1755
      : : : : :
Db      414  AspAspAlaThrValThrArgAsnAlaValThrLeuAlaLysGlyAlaThrIleLysLeu 433
      : : : : :
Qy      1756 TTAAGAGCTAAGCAGGCGCATCAGATCTCTTTAATGATCCCATCGAGATGGCAACGGA 1815
      : : : : :
Db      434  LeuAlaAlaSerGlyAspHleLysLeuCysPheTyAspProIleValThrLeuPro 453
      : : : : :
Qy      1816 AATAACCCAGCGCAGCTCTTCAAACTTCTAAATTAACGATGGTGA----- 1866
      : : : : :
Db      454  GluThrAlaProThr---AsnAspLysThrLeuThrIleAsnGlnAspLysThrSerSer 472
      : : : : :
Qy      1867 -----GGATACACAGGGGATATTGTTTT----- 1890
      : : : : :
Db      473  ThrProPheThrAsnTyIleGlyThrLeuLeuPheSerGlyAlaTyValAspSerGln 492
      : : : : :
Qy      1891 -----GCTAATGGAGCGATCTTTGTACCAAAATGTTACGATAGAGCAA 1935
      : : : : :
Db      493  SerAlaSerThrThrAlaAsnPheGluSerThrIleTyGlnLysValIleLeuGlyGly 512
      : : : : :
Qy      1936 GGAAGGATGTTCTTCGTTGAAAGGCAAAATATCATCGATGTAATCTCTAAGTCAG---ACA 1992
      : : : : :

```

```

Db      513  GlyLysLeuValLeuAlaAspLysAlaSerLeuSerValAlaSerPheThrGlnGluThr 532
      : : : : :
Qy      1993 GGTGGGAGTCTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACACACCACA 2052
      : : : : :
Db      533  AspSerIleLeuLeuMetAspAsnGlyThrThr-----LeuAlaIleThrGluHisSer 550
      : : : : :
Qy      2053 CAACAGCTCTCTGCGCT----- 2070
      : : : : :
Db      551  HisGlnThrProAlaAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 570
      : : : : :
Qy      2071 -----AATCAGTTGATCAGCTTTCAATCTGCAATTTGCTCTCTCTCTCTCTCTCTCT 2124
      : : : : :
Db      571  AsnThrAspGlyValIleSerIleAsnLeuHisValAsnIleSerSerLeuThrGlu 590
      : : : : :
Qy      2125 AACATCAGTTCAGAAATCTCTACCAATCTCTCCAGGCAAGATTCATCTCTGCAGTC 2184
      : : : : :
Db      591  GlnGlyGluGlyAlaLysLeuGluThr----- 599
      : : : : :
Qy      2185 ATTGGTAGCACCACTGCTGTTCTGTACAAATAGTGGGCTATCTTTTTCAGGATTTG 2244
      : : : : :
Db      600  -----LysAsnThrAspGlyThrIleThrLeuThrGlyHisValSerLeuAspAspVal 617
      : : : : :
Qy      2245 GATGATCAGCTTATGATAGTATGCTAGTGGCTAGGTTCTAATCAA---AAATCAATGTC 2301
      : : : : :
Db      618  SerGlyThrAlaTyGluAsnHisAspLeuPheAsnLysAspThrValThrIleAsnLeu 637
      : : : : :
Qy      2302 CTGAATTTACAGTTAGGG-----ACTAAGCCCCAGCTAATGCCCATCAGATTTGACT 2355
      : : : : :
Db      638  LeuSerLeuSerThrAlaGlyAspSerLysThrThrIleAsnGly---LeuAspLeuThr 656
      : : : : :
Qy      2356 CTA---GGGAATGAGTGCCTAAGTATGCTATCAAGAGCTGGAAGCTGCGTGGAT 2412
      : : : : :
Db      657  LeuArgGlyAspAlaGluProGlnTyGlyTyGlnGlySerTrpGlnLeuAlaIleTrpGlu 676
      : : : : :
Qy      2413 CCTAATACAGCAATAATGCTCTTATATCTCTGAAGCTACATGAGCTAAACCTGGGTAT 2472
      : : : : :
Db      677  AsnGlyAlaAspAlaAsnLysGlnLysIleLeuLysAlaThrTrpThrLysThrGlyPhe 696
      : : : : :
Qy      2473 AATCTCGGCTGAGCGAGTAGCTTCTTGGTTCCAAATAGTTTATGGGATCCATTATA 2532
      : : : : :
Db      697  ThrProAsnProGluArgGlnAlaSerLeuValProAsnSerLeuTrpGlyAlaPheIle 716
      : : : : :
Qy      2533 GATATACGATCTGGCATTCCAGCAATTCAGCAAGTGTGATGGCGCTCTTATGTCGA 2592
      : : : : :
Db      717  AspLeuSerMetAsnAlaLeuAlaThrAlaSerCysAspGlyPheGlyTyGlyLys 736
      : : : : :
Qy      2593 GGATATGGGTTCTGGAGTTTCGAATTTCTCTATCATCCGCGATGCTTTAGGTCAG 2652
      : : : : :
Db      737  GlyLeuTrpValAlaGlyIleSerAsnIlePheHisAspArgAsnSerValSerHis 756
      : : : : :
Qy      2653 GGATATCGGTATATTAGTGGGGTTATTCTCTAGGAGCAAACTCTTAC---TTTGGATCA 2709
      : : : : :
Db      757  GlyPheArgArgIleSerGlyGlyTyValIleGlyAlaAsnSerGlnThrValThrAsp 776
      : : : : :
Qy      2710 TCGATGTTTGGTTCAGCATTTCCGAAGTATTTGGTAGATCTAAAGATTATGATGTGT 2769
      : : : : :
Db      777  SerValPheGlyValAlaPheSerGlnIlePheAlaLysSerLysAspTyValValSer 796
      : : : : :
Qy      2770 CGTTCCCAATCATCTTCATAGGATCGTTTATCTATCTATCCCAACAGCTTTATGT 2829
      : : : : :
Db      797  SerAlaLysSerGlnAlaIleAlaGlySerAlaTyLeuSerValLysArgGlnLeuSer 816
      : : : : :
Qy      2830 GGATCTCTATTGTTTCGGAGATGGTATTATCTCTAGCTACCGGTTTGGGAATCAGCAT 2889
      : : : : :
Db      817  AsnThr---IlePheSerSerPheAlaAlaArgIleAsnTySerHisThrAsnGluAsp 835
      : : : : :
Qy      2890 ATGAAAACCTCATATACATTTGTCAGAGGAGGAGCATGTTCTGTTGGGATATAATCTGCTG 2949
      : : : : :
Db      836  MetLysThrArgTyThrPheIleProGluLysAspGlyAsnTrpAspAsnAsnCysTrp 855
      : : : : :
Qy      2950 GCTGGAGAGATTCGAGCGGATTCACGATGCTGATTAATCTCATCTAGCTCTATTTCAAT 3009
      : : : : :
Db      856  LeuGlyGluIleGlySerLeuProIleValLeuGlnIleThrLysLeuHisLeuAsn 875
      : : : : :

```

```

QY 3010 GAGTTGGCTCTTCTGCTCAAGCTGAGTTTCTTATGCGGATCATCAATCTTTTACAGAG 3069
DB 876 GlnIleIleProPheMetAsnValGlnLeuGlyTyrAlaGluHisGlySerPheLysGlu 895
QY 3070 GAAGGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCCTAAATCTATCACTTCTGTT 3129
DB 896 LysLeuAlaGluAlaArgSerPheCysSerArgLeuLeuAsnLeuAlaValProVal 915
QY 3130 GAGTGAAGTTTGATCGATGTTCTAGTACATCTCTTAATTAATATAGCTTTTGGCGCT 3189
DB 916 GlyPheLysIleAspArgSerHisSerHisProAspPheTyrSerLeuAlaIleSer 935
QY 3190 TATATCTGATGCTTATCGACCATCTCTGCTGCTAGTACGACACGCTCTCTATCCCATCA 3249
DB 936 TyrIleProAspValTyrArgAsnProGlyCysAsnThrLeuLeuLeuAlaAsnGly 955
QY 3250 GAGACATGACAAACAGATGCTTTTCAATTTAGCAAGACATGAGTGTGTTAGAGATCT 3309
DB 956 ValArgTyrLysThrProAlaThrAsnLeuAsnArgHisGlyLeuLeuMetGlnGlySer 975
QY 3310 ATGTATGCTCTCTCAAGATTAATATAGATATATGAGTATATGGCCATGAGATATGATCGA 3369
DB 976 ThrHisThrAlaValLeuSerAsnIleGluIlePheSerHisGlySerCysGluLeuArg 995
QY 3370 GATGCTTCTCGAGCTATGTTTGTAGTGCAGGAGTAGTCCGGTTC 3417
DB 996 SerSerArgAsnTyrAsnIleAsnValGlySerLysIleArgPhe 1011

```

## RESULT 4

Q5L6J7\_CHLAB

ID Q5L6J7\_CHLAB PRELIMINARY; PRT; 1024 AA.

AC Q5L6J7;

DT 01-FEB-2005 (TrEMBLrel. 29, Created)

DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Polymorphic outer membrane protein.

GN Name:pmprG; Ordered locus names: CAB269;

OS Chlamydomophila abortus.

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.

OX NCBI\_TaxID=83555;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=S2673;

RX PubMed=15837807; DOI=10.1101/gr.3684805;

RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,

RA Livingstone M., Cerdono-Tarraga A.-M., Harris B., Doggett J.,

RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,

RA Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.

RT "The Chlamydomophila abortus genome sequence reveals an array of

RT variable proteins that contribute to interspecies variation.";

RL Genome Res. 15:629-640(2005).

DR EMBL; CR848038; CAH63725.1; -; Genomic\_DNA.

KW Complete proteome.

SQ SEQUENCE 1024 AA; 108663 MW; 268E61415C8FD434 CRC64;

## Alignment Scores:

```

Pred. No.: 3.43e-109 Length: 1024
Score: 1777.00 Matches: 415
Percent Similarity: 53.5% Conservative: 169
Best Local Similarity: 38.0% Mismatches: 362
Query Match: 22.5% Indels: 146
DB: 2 Gaps: 28

```

US-10-701-844-1 (1-4435) x Q5L6J7\_CHLAB (1-1024)

QY 382 ATGCAACGCTTTCCATAGATCTTCTTCAATGATTCAGTATTCTTGTGCTCT 441

DB 1 MetLysAlaSerPheArgLysPheLeuValSerThrThrLeuThrLeuProCys- 18

QY 442 TTAATGGGGGGGATATGACGAGAAATCATGTCTCAAGAAATTTACGATGGGAG 501

DB 19 ---SerPheGlnAlaPheSerLeuGluLeuValProAsnGlyThrTyrAsnGlyAsp 37

```

QY 502 ACGTTAACTGTATCATTTCCCTTACTCTTATAGGAGATCCGAGTGGAGTACTACTCTTTT 561
DB 38 ---LeuArgGluMetPheProTyrThrIleThrSerAsnProGluGlyThrThrAlaIle 56
QY 562 TCTGAGGAGAGTTAAATTAATAATCTTGACAAATCTTATTTGACAGCTTTGCTTTAAGT 621
DB 57 LeuSerGlyAsnLeuAsnIleLeuAsnLeuAspAsnSerMetAlaAlaThrProSerSer 76
QY 622 TGTTTTGGGAACTTATTAGGAGTTTACTGTTTTAGGAGAGGACACTCGTTGTGACTTTC 681
DB 77 CysPhePheAsnSerAlaGlySerMetThrIleIleGlyLysGlyHisAspValThrPhe 96
QY 682 GAGAACATACGACATCTTACAAATAGGGCAGCTCTAAGT-----AATAGCGCT 729
DB 97 ThrAsnLeuArgThrSerValAsnGlyAlaAlaLeuSerSerLeuIleThrSerSerPro 116
QY 730 GCTGATGAGCTGTTTACTATTGAGGGTTTTAAAGAATTTATCTTTTCCATTTGCAATTC 789
DB 117 GluSerPheSerTyrThrIleThrGlyValLysThrPheSerCysSerAsnCysSerAla 136
QY 790 TTACTTGGCGTACTGCTCTCAACGACTAATAAGGGTAGCCAGACTCCGACGACAA 849
DB 137 LeuLeu-----GlyArgAsnAsnSerAsnSerIle 146
QY 850 TCTACACCGCTCTAATGGTACTATTATTCTTAAACACAGATCTTTTGTACTCAATATGAG 909
DB 147 LeuSerProLysGlyGlyAlaValTyrSerLysSerProIlePhePheLysAsnIleGln 166
QY 910 AAGTTCTCATCTATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGC 969
DB 167 AsnLeuIlePheLysAspAsnCysAlaAlaAspAsnGlyGlyAlaLeuTrpGlyGlnVal 186
QY 970 TTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGAT 1029
DB 187 ValAspIleSerAsnValThrLysSerLeuLysPheLeuSerAsnValGly---AlaAsn 205
QY 1030 GGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCTAAACGAGGCTCT---ATT 1086
DB 206 GlyGlyAlaIleGlyAlaSerThrSerLeuAsn-----ValThrArgCysProSerIle 223
QY 1087 GCCTTTGTAGCGAATTTGCGAGGATGTAAGAGGGGAGGAGTCTGCTGTTCCAGGATGGG 1146
DB 224 LeuPheArgSerAsnSerAlaSerArgLeuGlyGlyAlaIle----- 237
QY 1147 CAGCAGGAGTGTCACTATCTACTTCAACAGAGATCCA----- 1185
DB 238 ---HisSerValAsnProGlnThrProProProProProProGlyAsnGlyValIle 256
QY 1186 -----GTAGTAAGTTTTCAGAAATCTGCGGTAGAGTTTGTATGGGAACGTAGCCCGA 1239
DB 257 AsnThrValValAsnPheSerAspAsnGlySerValGlnPheAspSerAsnAsnAlaLys 276
QY 1240 GTAGGAGGAGGATTTACTCTCTACCGGAACGTTGCTTTCTGTAATATGAAACAACTTG 1299
DB 277 SerGlyGlyAlaIleTyrSerLysGlyAsnIleAsnPheSerAsnValGlnLeuLeu 296
QY 1300 TTTCTCAACATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1359
DB 297 MetGlnAsnAsnThrAlaSerPro-----Glu 305
QY 1360 GCTTCTTAATACGAGTAATAATTACGGAGATGGAGGAGCTATCTTCTGTAAGAATGGTGC 1419
DB 306 IleGlyAspSerAsnGluValLeuGlyGlnGlyAlaIleTyrCys-----Thr 322
QY 1420 CAAGCAGGATCCCAATAACTCTGATCAGTTTCTCTTGTATGGAGGAGGAGTA----- 1470
DB 323 GlnAlaThrAlaThrLysLysAlaAlaProValPheThrGlyLeuThrIleThrAsnGln 342
QY 1471 -----GTTTCTTTTAGTACCAATGATGCTGCGGGAAGGGGAGCTATTTATGCCAAA 1524
DB 343 HisAspIlePhePheAsnAsnPheAlaAsnAlaGlyGlyAlaIleTyrGlyGlu 362

```

Qy	1525	AAGCTCTCGGTGCTTAACATGTGGCCCTGTACAATATTTTAAAGGAATATCGCTAATGATGGT	1581
Db	363	LySValSerIleThrSerSerGlyArgThrValPheThrAenAenThrAlaLysThrGly	382
Qy	1585	GGAGCGATTTATTTAGAGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATT	1644
Db	383	GlyAlaIleTyrIleAlaAspAenSngIgluLeuSerLeuSerAlaAspTyrGlyAenMet	402
Qy	1645	ATTTTTCGATCGGAATCTTAAAGAACAGCCAAAGAGAATGCTGCCGATTTAATGGCGTA	1704
Db	403	ThrPheTyrAenAenLeu-----AenThrAenSngIgly	413
Qy	1705	ACTGTGTCCTACAAGCCATTTTCGATGGATCGGAGGAGAAATAACAGCATTAAGAGCT	1764
Db	414	SerProLysArgAenAlaValThrLeuGlyLysGlyAlaThrIleLysLeuAlaAa	433
Qy	1765	AAAGCAGGCGATCAGATCTCTTTAATGATCCCATCGAGATGGCAACGGAAATAACCAG	1824
Db	434	SerGlyAspHisLysLeuCyPheTyrAspProIleValThrThrLeuProGluLysAa	453
Qy	1825	CCAGCGCAGCTCTCCAAACTCTTAAAAATAAAGCATGTCGAAGA-----	1869
Db	454	Pro----AenGlySerAenThrLeuThrIleAenProAspArgGlyAenSerAlaProPhe	472
Qy	1870	-----TACACAGCGGATATTGTTTTT-----GCTAAT	1896
Db	473	ThrAenTyrIleGlyThrValLeuPheSerGlyAlaHisAlaAenAlaGluSerAlaPro	492
Qy	1897	GGAAGCAGTACTTTGTACCAAAATGTTACGATAGACGACGAGAGGATGTTCTTCGTGAA	1956
Db	493	AsnAlaSerThrIleTyrGlnLysValIleLeuGlyGlyLysLeuValLeuAlaAsp	512
Qy	1957	AAGCCAAAATATCAGTGAATTCCTCAAGTCAGACAGGTGGAGT---CNGTATATGGAA	2013
Db	513	LysAlaSerLeuSerValValSerPheAspGlnGlnSerAspSerIleLeuLeuMetAsp	532
Qy	2014	GCTGGGAGTACATG-----	2027
Db	533	AsnGlyThrSerLeuThrIleThrGluHisSerTyrAlaThrSerSerGluAlaValAla	552
Qy	2028	-----GGATTTTGTAACCTCCACACACC	2051
Db	553	AlaAlaAlaSerValValThrAlaAlaGluValSerSerValThrSerSerThrAlaAa	572
Qy	2052	ACAACAGCTCTCGCGCTAATCAG-----TTGATCACGCTTCTCCATCT	2096
Db	573	ThrAlaAlaSer-ValThrAenGlnAlaAenSerAspGlyValSerIleLysAeLe	592
Qy	2097	GCATTTGCTCTTCTTCTTTGTTAGCAAACAATGCAGTTACGAATCTCTCTACCAATCC	2156
Db	592	uHisIleAenLeuGlySerLeu-----	599
Qy	2157	TCCAGCCGCAAGATTCATCTCTCGAGTCATTTGGTACG---ACAACCTGCTGGTTCGTATC	2213
Db	600	----ThrGlnAspGlyGluGlyAlaLysIleGluThrLysAenThrSerGlyThrIleTh	618
Qy	2214	AATTAGTGGGCCCTCTTTTTCAGGATTTGGATGATACAGCTTATCATAGTAGTATGATTG	2273
Db	618	IleSerGluHisIleSerLeuAspValSerGlyAsnAlaTyrGluAenHisAspIle	638
Qy	2274	GCTAGGTTCTAATCMAAAATCAATGCTCGAAATTTACAGTTAGGAGACTAAGCCCCAGC	2333
Db	638	ePheAenArgAen---ThrValThrLeuLysValLeuSerLeuSerThrAlaGlyAspAs	657
Qy	2334	TAAT---GCCCCATCAGATTTGACTCTA-----GGGAATGAGATGCTAAGTATGG	2381
Db	657	nLysIleThrGluSerAspLeuGlnLeuThrProArgGlyAspAlaAspProGlnTyrGly	677
Qy	2382	CTATCAGGGAAGCTGGAGCTTCGGTGGGATCCTAATACAGCAAAATAATCGT-----	2433
Db	677	YtyrGlnGlySerTyrLysLeuSerIleTyrGluAenGlyThr-----AsnGlyAspAlaGly	695
Qy	2434	-----CCTTATACTCTGAAGCTACATGGACTGAAACTGGGTATAATCTCGGGCTCAGCG	2489

RESULT 5
PMPH_CHL'
ID PMP
AC 084
DT 16-
DT 16-

DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Probable outer membrane protein pmph precursor (Polymorphic membrane  
 DE protein H).  
 GN Name=pmph; OrderedLocusNames=C7872;  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=D/UW-3/Cx;  
 RX MEDLINE=9900809; PubMed=9784136; DOI=10.1126/science.282.5389.754;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)  
 CC (Potential).  
 CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; AR001360; AAC68470.1; -; Genomic\_DNA.  
 DR PIR; H71460; H71460.  
 DR PHCI-2DPAGE; O84880; -.  
 DR InterPro; IPR005546; Auto\_transptbeta.  
 DR InterPro; IPR011427; ChlamPMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 6.  
 DR Pfam; PF07548; ChlamPMP\_M; 1.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 4.  
 DR Complete proteome; Membrane; Multigene family; Outer membrane; Signal.  
 KW SIGNAL 1 24 Potential.  
 FT CHAIN 25 1016 Probable outer membrane protein pmph.  
 FT SEQUENCE 1016 AA; 107905 MW; E691912C3A2BD6F7 CRC64;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.05e-96 Length: 1016  
 Score: 1589.00 Matches: 311  
 Percent Similarity: 92.7% Conservatives: 5  
 Best Local Similarity: 91.2% Mismatches: 10  
 Query Match: 20.2% Indels: 15  
 DB: 1 Gaps: 2  
 US-10-701-844-1 (1-4435) x PMPH\_CHLTR (1-1016)  
 QY 3451 ATGCTTTTCTTTGAGATCTACATCATTTTGTGTTTGTGTTCTCTATTCG 3510  
 Db 1 MetPhosphorylaseSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrrSer 20  
 QY 3511 TATGGATTCGCGAGCTCTCTCAAGTGTGTTAAACGCTTAATGACCACTCTCTTTAAGGGA 3570  
 Db 21 TyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrProPheLysGly 40  
 QY 3571 GAGCATGTTTACTGTAATGAGACTGCGCTTTTGTCTCATGCTCTATGCGAGGCTGAGAA 3630  
 Db 41 AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn 60  
 QY 3631 GGTTCGATTATCTCAGCTAATGCGCAATTTTACGATTACCGGACAAACCATACATTA 3690  
 Db 61 GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu 80  
 QY 3691 TCATTTACAGATTCCTCAAGGCCAGTTCTCTCAAAATTTATGCTTCATTTACGAGGAGAG 3750  
 Db 81 SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu 100  
 QY 3751 ACATTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCTTCGCGGAGAA 3810

Db 101 ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu 120  
 QY 3811 AAGGGATGATCTCCGGGAAAACCGTAGTATTTCCGAGCAGGCGAAGTGCATTTCTGG 3870  
 Db 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrp 140  
 QY 3871 GATAACTCCGCGGGTATTTCTCTTATCTACTGTGCCAACCTCATCATCACTCCGCT 3930  
 Db 141 AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro 160  
 QY 3931 GCT-----CCACAGTTAGTATGCTCGG 3954  
 Db 161 AlaProAlaProAlaProAlaSerSerSerLeuSerProThrValSerAspAlaArg 180  
 QY 3955 AAAAGGCTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAAAAAGGGTC 4014  
 Db 181 LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLysGlyVal 200  
 QY 4015 ATGTTCCGATAATAATCCGGGAATTTCCGAACAGTTTTTCGAGGTAAAGATAATAATAAT 4074  
 Db 201 MetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsn 220  
 QY 4075 GCTGCTGGTGGAGGCGAGTGGTTCGCTACACCATCAAGTACGACTTTTACAGTTAAAAA 4134  
 Db 221 AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLysAs 238  
 QY 4135 CTGTAAAGGAAAAGTTTCTTTCACAGATAACGTAGCTCTTCGCGGCGGAGTGGTTTA 4194  
 Db 238 nCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyValValTy 258  
 QY 4195 TAAAGCATTTGTCCTTTTCAAGACAATGAAGGAGGCATATTCTTCCGAGGGAACACAGC 4254  
 Db 258 rLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePheArgGlyAsnThrAl 278  
 QY 4255 ATACGATGATTTAAGGATTTCTGCTACTAATCAGATCAGATACGAGACGAGGAGG 4314  
 Db 278 atyAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyG 298  
 QY 4315 CGGTGGAGGAGTATTTGCTCTCCAGATGATTTCTGTAAGGCAATAAAGGTTTC 4374  
 Db 298 yGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySe 318  
 QY 4375 TATTGTTTTTGATTAACAATTTGCAAAAGCGAGAGCGCAAGCATCTTAACGAAAGATT 4434  
 Db 318 rIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPh 338  
 QY 4435 C 4435  
 Db 338 e 338  
 RESULT 6  
 Q84FUT7 CHLTR PRELIMINARY; PRT; 981 AA.  
 AC Q84FUT7;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polymorphic membrane protein H (Fragment).  
 GN Name=pmph;  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22483673; PubMed=12595433;  
 RX DOI=10.1128/IAI.71.3.1200-1208.2003;  
 RX Stothard D.R., Toth G.A., Batteiger B.E.;  
 RA "Polymorphic membrane protein H has evolved in parallel with the three  
 RA disease-causing groups of Chlamydia trachomatis.";  
 RL Infect. Immun. 71:1200-1208(2003).  
 DR EMBL; AV184168; AAC030032.1; -; Genomic\_DNA.  
 DR GO; GO:0019867; C:outer membrane; IEA.



DR InterPro; IPR005546; Auto.transp.bet.  
 DR InterPro; IPR011427; Chlam\_PMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF07548; Chlam\_PMP\_M; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 6.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 4.  
 DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
 FT NON TER 1  
 FT NON TER 981  
 SQ SEQUENCE 981 AA; 104559 MW; 83C957685E57EC10 CRC64;

Alignment Scores:  
 Pred. No.: 3.51e-95 Length: 981  
 Score: 1566.00 Matches: 303  
 Percent Similarity: 99.7% Conservative: 0  
 Best Local Similarity: 99.7% Mismatches: 0  
 Query Match: 19.9% Indels: 1  
 DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x Q84F07\_CHLTR (1-981)

QY 3526 TCTCTCAAGTGTAAACGCTTAATGTACCACTCTTTTAAGGAGAGAGATGTTTACTTG 3585  
 DB 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValTyrlLeu 20

QY 3586 AATGGAGACTCGCTTTTGTCTAATGTCTATCGAGAGCTGAAGAAGGTTCGATTATCTCA 3645  
 DB 21 AsnGlyAspCysAlaPheValAsnValTyrlAlaGlyAlaGluGlySerIleIleSer 40

QY 3646 GCTAATGGCGACAATTAAACGATTACCGGACAAACCATATCATTTACAGATTCT 3705  
 DB 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60

QY 3706 CAAGGGCCAGTCTTCAAAATGTATGCTTCACTTCACTGAGGAGAGACACTTACTCTGAGA 3765  
 DB 61 GlnGlyProValLeuGlnAsnTyrlAlaPheIleIleSerAlaGlyGluThrLeuThrLeuArg 80

QY 3766 GATTTTTCGAGTCTGATGTTCTCGAAAATGTCTTTCGCGAGAGAAAGGAATGATCTCC 3825  
 DB 81 AspPheSerSerLeuMetPheSerIleAsnValSerCysGlyGluLysGlyMetIleSer 100

QY 3826 GGGAAACCGTAGTATTTCCGGAGCGGCGAAGTAGATTTTCGGGATACTCCGTCGGGG 3885  
 DB 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrPaspAsnSerValGly 120

QY 3886 TATTCTCTTTTATCTACTGTCCCAACCTCATCATCACTCCGCTGCTCCCAACAGTTAGT 3945  
 DB 121 TyrSerProLeuSerThrValProThrSerSerSerThrProProAlaProThrValSer 140

QY 3946 GATGCTCGAAAGGTCATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAA 4005  
 DB 141 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 160

QY 4006 AAAGGGTCATGTCGATATAATAATCCCGGAATTTTCGGAACAGTTTTCGAGGTAAGAT 4065  
 DB 161 LysGlyValMetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyLysAsn 180

QY 4066 AATAATATGCTGTGTGGAGGAGTGTTCCGCTACACCATCAAGTACGACTTTTAC 4125  
 DB 181 AsnAsnAsnAlaGlyGlyGlyGlySer-GlySerAlaThrProSerSerThrThrPheTh 200

QY 4126 AGTTAAAACTGTAAAGGAAAGTTTCTTTTCCAGAGATAACGATAGCTCTTCGCGAGCGG 4185  
 DB 200 rValIleAsnCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyGly 220

QY 4186 AGTGTTTATAAAGGCATGTGCTTTTCAAGACAAATGAAGAGGCATATTTCTTCGAGG 4245  
 DB 220 yValValTyrlLysGlyIleValLeuPheLysAspAsnGluGlyIlePhePheArgGly 240

QY 4246 GAACACAGCATACCATGATTTAAGGATTTCTGTGCTACTTAATCAGGATCAGATACGGA 4305

DB 240 yAsnThrAlaTyrlAspAspLeuArgIleLeuAlaAlaThrAsnGlnAspGlnAsnThrGl 260  
 QY 4306 GACAGAGCGGTGGAGAGTATTGCTCTCCAGATGATCTGTAAAGTTTGAAGCAA 4365  
 DB 260 uThrGlyGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAs 280  
 QY 4366 TAAAGGTTCTATTGTTTGTGATTACACTTTGCAAAAGGCGAGGCGAGACATCCTAAC 4425  
 DB 280 nLysGlySerIleValPheAspTyrlAsnPheAlaLysGlyArgGlyGlySerIleLeuTh 300

QY 4426 GAAAGAATTC 4435  
 DB 300 rLysGluPhe 303

RESULT 7  
 Q84F06\_CHLTR  
 ID Q84F06\_CHLTR PRELIMINARY; PRT; 980 AA.  
 AC Q84F06;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polymorphic membrane protein H (Fragment).  
 GN Name=bmpH;  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN NUCLEOTIDE SEQUENCE.  
 RP MEDLINE=22483673; PubMed=12595433;  
 RX DOI=10.1128/IAI.71.3.1200-1208.2003;  
 RA Stothard D.R., Toth G.A., Basteiger B.E.;  
 RT "Polymorphic membrane protein H has evolved in parallel with the three  
 disease-causing groups of Chlamydia trachomatis.";  
 RL Infect. Immun. 71:1200-1208(2003).  
 DR EMBL; AV184169; AAC03033.1; -; Genomic\_DNA.  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR InterPro; IPR005546; Auto.transp.bet.  
 DR InterPro; IPR011427; Chlam\_PMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF07548; Chlam\_PMP\_M; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 6.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 4.  
 DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
 FT NON TER 1  
 FT NON TER 980  
 SQ SEQUENCE 980 AA; 104432 MW; BC9B7E33BBD04233 CRC64;

Alignment Scores:  
 Pred. No.: 3.22e-94 Length: 980  
 Score: 1551.50 Matches: 302  
 Percent Similarity: 99.3% Conservative: 0  
 Best Local Similarity: 99.3% Mismatches: 0  
 Query Match: 19.7% Indels: 2  
 DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x Q84F06\_CHLTR (1-980)

QY 3526 TCTCTCAAGTGTAAACGCTTAATGTACCACTCTTTTAAGGAGAGAGATGTTTACTTG 3585  
 DB 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValTyrlLeu 20

QY 3586 AATGGAGACTCGCTTTTGTCTAATGTCTATCGAGAGCTGAAGAAGGTTCGATTATCTCA 3645  
 DB 21 AsnGlyAspCysAlaPheValAsnValTyrlAlaGlyAlaGluGlySerIleIleSer 40

QY 3646 GCTAATGGCGACAATTAAACGATTACCGGACAAACCATATCATTTACAGATTCT 3705  
 DB 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60

QY 3706 CAAGGGCCAGTCTTCAAAATGTATGCTTCACTTCACTGAGGAGAGACACTTACTCTGAGA 3765

Db 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuArg 80  
QY 3766 GATTTTCGAGTCTGATGTTTCGAAAATGTTTCTCGGAGAAAAGGGAATGATCTCC 3825  
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluGlyMetIleSer 100  
QY 3826 GGGAAAACCGTCAGTATTCCGAGCAGCGAAGTATTCTCGGATAAATCTCCGTGGGG 3885  
Db 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrPheAsnSerValGly 120  
QY 3886 TATTTCTCCTTTATCTACTGTGCCACCTCATCATCACTCCGCTGCTCCAAAGTTAGT 3945  
Db 121 TyrSerProLeuSerThrValProThrSerSerSerThrProProAlaProThrValSer 140  
QY 3946 GATGCTCGGAAGGTCATATTTTCTGTAGAGACTAGTTTTCGAGATCTCAGCGCTCAA 4005  
Db 141 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 160  
QY 4006 AAAGGGCTCATGTCGATAATATGCGGGGAATTTCCGGAACAGTTTTCGAGTAAAGAT 4065  
Db 161 LysGlyValMetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyLysAsn 180  
QY 4066 AATAATAATGCTGGTGGAGCGAGTGGGTTCCGCTACACCATCAAGTACGACTTTTAC 4125  
Db 181 AsnAsnAsnAlaGlyGlySerGlyGlySerGlySerAlaThrProSerSerThrPheTh 200  
QY 4126 AGTTAAATCTGTAAGGGAAGTGTCTTCAAGACAATGAAGGAGGATATCTCCGAGG 4185  
Db 200 rValLysAsnCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGly 220  
QY 4186 AGTGGTTTATAAGGCAATGTCGTTTCAAGACAATGAAGGAGGATATCTCCGAGG 4245  
Db 220 yValValTyrLysGlyIleValLeuPheLysAspAsnGluGlyGlyIlePhePheArg 240  
QY 4246 GAACACAGCATACGATGATTAAGGATTTCTGCTACTACTAATCAGGATCAGAAACGA 4305  
Db 240 yAsnThrAlaTyrAspAspLeuArgIleLeu---AlaThrAsnGlnAsnThrG 259  
QY 4306 GACAGAGCGGPGAGAGTATTTGCTCTCAGATGATCTGTAAGTTTGAAGGCAA 4365  
Db 259 uthrGlyGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAs 279  
QY 4366 TAAAGTCTTATGTTTGTATTACACTTGCATAAAGGAGGAGGAGGAGGATCTTAAC 4425  
Db 279 nLysGlySerIleValPheAspTyrAsnPheAlaLysGlyArgGlySerIleLeuTh 299  
QY 4426 GAAAGAAATC 4435  
Db 299 rLysGluPhe 302

RESULT 8  
Q84F08 CHLTR  
ID Q84F08 CHLTR PRELIMINARY; PRT; 980 AA.  
AC Q84F08  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Polymorphic membrane protein H (Fragment).  
GN Name=pmph;  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22483673; PubMed=12595433;  
RA DOI=10.1128/IAI.71.3.1200-1208.2003;  
RA Stothard D.R., Toth G.A., Basteiger B.E.;  
RT "Polymorphic membrane protein H has evolved in parallel with the three  
RT disease-causing groups of Chlamydia trachomatis";  
RL Infect. Immun. 71:1200-1208(2003).  
RL EMBL; AY184167; AAC03031.1; -; Genomic\_DNA.  
DR GO; GO:0019867; C:outer membrane; IEA.  
DR InterPro; IPR005546; Auto\_transpbeta.

DR InterPro; IPR011427; ChlamPMP\_M.  
DR InterPro; IPR003368; Chlamydia\_PMP.  
DR InterPro; IPR002016; Peroxidase.  
DR Pfam; PF03797; Autotransporter; 1.  
DR Pfam; PF07548; ChlamPMP\_M; 1.  
DR Pfam; PF02415; Chlam\_PMP; 6.  
DR TIGRFAMs; TIGR01376; POMP\_repeat; 4.  
DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
FT NON\_TER 1 980  
SQ SEQUENCE 980 AA; 104460 MW; 27PF6C8EA9F43B20P CRC64;  
  
Alignment Scores:  
Pred. No.: 3,22e-94 Length: 980  
Score: 1551.50 Matches: 302  
Percent Similarity: 99.3% Conservative: 0  
Best Local Similarity: 99.3% Mismatches: 0  
Query Match: 120 Indels: 2  
DB: 2 Gaps: 1  
  
US-10-701-844-1 (1-4435) x Q84F08\_CHLTR (1-980)  
QY 3526 TCTCTCAAGTGTAAACGCTTAATGTAACCTCTCTTTAAGGAGACGATGTTTACTTG 3585  
Db 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValTyrLeu 20  
QY 3586 AATGAGAGCTGCGCTTTTGTCAATGCTATGAGAGAGCTGAAAGAGGTTTCGATTATCTCA 3645  
Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluGluGlySerIleSer 40  
QY 3646 GCTAATGCGCAATTTAAAGTATACCGGACAAAACCATACATTATCATTTACAGATTCT 3705  
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60  
QY 3706 CAAGGCGCGATCTTCAAAATATGCTTCAATTCAGCAGAGAGACACTTACTCTCGAGA 3765  
Db 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuArg 80  
QY 3766 GATTTTTCGAGTCTGATGTTCTCGAAAATGTTTCTTGGGAGAAAAGGGAATGATCTCC 3825  
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100  
QY 3826 GGGAAAACCGTCAGTATTTCGAGCAGCGAAGTATTCTCGGATAAATCTCCGTGGGG 3885  
Db 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrPheAsnSerValGly 120  
QY 3886 TATTTCTCCTTTATCTACTGTGCCACCTCATCATCACTCCGCTGCTCCAAAGTTAGT 3945  
Db 121 TyrSerProLeuSerThrValProThrSerSerSerThrProProAlaProThrValSer 140  
QY 3946 GATGCTCGGAAGGTCATATTTTCTGTAGAGACTAGTTTTCGAGATCTCAGCGCTCAA 4005  
Db 141 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 160  
QY 4006 AAAGGGCTCATGTTTCGATAATATGCGGGAATTTCCGGAACAGTTTTCGAGTAAAGAT 4065  
Db 161 LysGlyValMetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyLysAsn 180  
QY 4066 AATAATAATGCTGGTGGAGCGAGTGGGTTCCGCTACACCATCAAGTACGACTTTTAC 4125  
Db 181 AsnAsnAsnAlaGlyGlyGlySerGlySerAlaThrProSerSerThrThrPheTh 200  
QY 4126 AGTTAAATCTGTAAGGGAAGTGTCTTCAAGACAATGAAGGAGGATATCTCCGAGG 4185  
Db 200 rValLysAsnCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGly 220  
QY 4186 AGTGGTTTATAAGGCAATGTCGTTTCAAGACAATGAAGGAGGATATCTCCGAGG 4245  
Db 220 yValValTyrLysGlyIleValLeuPheLysAspAsnGluGlyGlyIlePhePheArg 240  
QY 4246 GAACACAGCATACGATGATTAAGGATTTCTGCTACTACTAATCAGGATCAGAAACGA 4305  
Db 240 yAsnThrAlaTyrAspAspLeuArgIleLeu---AlaThrAsnGlnAsnThrG 259

```

Qy 4306 GACAGGAGCGGTGAGAGGATTTTGTCTCCAGATGATTCGTAAAGTTTGAAGGCAA 4365
Db 259 UTHrGlyGlyGlyGlyValIleCysSerProAspAspSerValIysPheGluGlyAs 279
Qy 4366 TAAAGGTTCTATTGTTTTGATTCAACTTTGCAAAAGCGAGAGCGGAAGCATCTTAAC 4425
Db 279 nLyGlySerIleValPheAspTyAsnPhAlaLysGlyArgGlyGlySerIleLeuTh 299
Qy 4426 GAAAGAATTC 4435
Db 299 rLyGluPhe 302

RESULT 9
ID Q83U76_CHLTR PRELIMINARY; PRT; 989 AA.
AC Q83U76
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Polymorphic membrane protein H (Fragment).
GN Name=pmpH;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22483673; PubMed=12595433;
RX DOI=10.1128/JAI.71.3.1200-1208.2003;
RA Stothard D.R., Toth G.A., Basteiger B.E.;
RT "Polymorphic membrane protein H has evolved in parallel with the three
RL disease-causing groups of Chlamydia trachomatis.";
RL Infect. Immun. 71:1200-1208(2003).
DR EMBL; AY184165; AA030029.1; -; Genomic DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; AutoTranspBeta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF04415; Chlam_PMP; 6.
DR TIGRFAMs; TIGR01376; POMP repeat; 4.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 989
SQ SEQUENCE 989 AA; 104934 MW; 75D5EAPF4633F691 CRC64;

Alignment Scores:
Pred. No.: 4.5e-88 Length: 989
Score: 1459.00 Matches: 287
Percent Similarity: 93.0% Conservative: 5
Best Local Similarity: 91.4% Mismatches: 9
Query Match: 18.5% Indels: 13
DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x Q83U76_CHLTR (1-989)
Qy 3526 TCTCTCAAGTGTAAAGCGCTTAAGTACCACTCTTTTAAAGGAGAGGATGTTTACTTG 3585
Db 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspValTyLeu 20
Qy 3586 AATGAGACTCGCTTTTGTCTAATGTCATCGAGAGCTGAAGAGGTTCCATTATCTCA 3645
Db 21 AsnGlyAspCysAlaPheValAsnValTyralaGlyAlaGluAsnGlySerIleIleSer 40
Qy 3646 GCTAATGCGGCAATTAAACGATTACCGGACAAACCATACATTATCATTTACAGATTCT 3705
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60
Qy 3706 CAAGGGCCAGTCTTCAAAATATGCTTCACTTTTCAGGAGGAGACACTTACTCTGGA 3765
Db 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspValTyLeu 20

```

```

Db 61 GlnGlyProValLeuGlnAsnTyAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80
Qy 3766 GATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCTTCGAGAAAAAGGAATGATCTCC 3825
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100
Qy 3826 GCGAAAAACCTGAGTATTTCGAGAGCGAGGCGAAGTGAATTTCTCGGATAACTCCGTCGGG 3885
Db 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrpAspAsnSerValGly 120
Qy 3886 TATTCTCTTTTATCTACTGTGCCAACCTCATCATCACTCCGCTGCT----- 3933
Db 121 TySerProLeuSerThrValProAlaSerThrProThrProProAlaProAlaProAla 140
Qy 3934 -----CCACAGTTTAGTGTGCTCGGAAAGGCTGCTATTTTCTGTGTA 3975
Db 141 AlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePheSerVal 160
Qy 3976 GAGCTAGTTTGGAGATCTCAGCGGTCAAAAAAGGGGTGATGTCGATATATATGCCGG 4035
Db 161 GluThrSerLeuGluIleSerGlyValLysGlyValMetPheAspAsnAsnAlaGly 180
Qy 4036 AATTTCCGAACAGTTTTTTCGAGGTAAAGATAATAAATCTGCTGCTGAGGCGAGTGGG 4095
Db 181 AsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGlyGlySer-Gl 200
Qy 4096 TTCGCTACACCATCAAGTACGACTTTTACAGTTTAAAAAACTGTAAGGGGAAGTTCTTT 4155
Db 200 YSerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyValSerPh 218
Qy 4156 CACAGATAAGTAGCTCTTCGCGAGCGGAGGTGTTTATAAGGCATGTCGCTTTTCAA 4215
Db 218 eThrAspAsnValAlaSerCysGlyGlyValValValTyLysGlyThrValLeuPheLy 238
Qy 4216 AGACAAATGAAGGAGGCATATTTCTTCGAGGGAACACAGCATACGATGATTTAAGGATTCT 4275
Db 238 sAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyAspAspLeuGlyIleLe 258
Qy 4276 TGCTGCTACTAATCAGGATCAGAAATACGAGACAGGAGCGGTGGAGAGGTTATTTGCTC 4335
Db 258 uAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyValIleCysSe 278
Qy 4336 TCCGAGATGATCTGTAAGTTTGAAGCAATAAAGTTCTATTGCTTTTGAATTAACAATT 4395
Db 278 rProAspAspSerValLysPheGluGlyAsnLysGlySerIleValPheAspTyAsnPh 298
Qy 4396 TGCAAAAGCGAGGCGGAGGATCCTTAACGAAAGATTC 4435
Db 298 eAlaLysGlyArgGlyGlySerIleLeuThrLysGluPhe 311

RESULT 10
Q84FU9_CHLTR
ID Q84FU9_CHLTR PRELIMINARY; PRT; 989 AA.
AC Q84FU9
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic membrane protein H (Fragment).
GN Name=pmpH;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22483673; PubMed=12595433;
RX DOI=10.1128/JAI.71.3.1200-1208.2003;
RA Stothard D.R., Toth G.A., Basteiger B.E.;
RT "Polymorphic membrane protein H has evolved in parallel with the three
RL disease-causing groups of Chlamydia trachomatis.";
RL Infect. Immun. 71:1200-1208(2003).
DR EMBL; AY184164; AA030028.1; -; Genomic DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; AutoTranspBeta.

```

DR InterPro; IPR011427; ChlamPMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF07548; ChlamPMP\_M; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 6.  
 DR TIGRfam; TIGR01376; POMP repeat; 4.  
 DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
 FT NON\_TER 989 989  
 FT NON\_TER 1  
 SQ SEQUENCE 989 AA; 104953 MW; 75D5F65D4F21E371 CRC64;

Alignment Scores:  
 Pred. No.: 4,5e-88 Length: 989  
 Score: 1459.00 Matches: 287  
 Percent Similarity: 93.0% Conservative: 5  
 Best Local Similarity: 91.4% Mismatches: 13  
 Query Match: 18.5% Indels: 9  
 DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x Q84FU9\_CHLTR (1-989)

QY 3526 TCTCTCAAGTGTAAACGCTAATGTAAACACTCTCTTTAAGGAGACGATCTTTACTTG 3585  
 Db 1 SerProGlnValLeuThrProAnValThrThrProPheLysGlyAspValTyrLeu 20

QY 3586 AATGGAGACTGCGCTTTCTCAATGTCATGAGGAGCTGAAGAGTTTCGATTATCTCA 3645  
 Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleSer 40

QY 3646 GCTAATGGGACAAATTACGATTACCGGACAAACCATACATTATCATTTACAGATTCT 3705  
 Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHieThrLeuSerPheThrAspSer 60

QY 3706 CAAGGCCAGTCTCTCAAAATATGCTCTTCATTTACAGAGGAGACACTTACTCTGAGA 3765  
 Db 61 GlnGlyProValLeuGlnAnThrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80

QY 3766 GATTTTCGAGTCTGATGTCGAAAAATGTTCTTCGGGAGAAAAAGGAATGATCTCC 3825  
 Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100

QY 3826 GGGAAACGTCGATTTTCGGGAGCAGGAGTGTATTTCTCGGATACCTCCGNGGG 3885  
 Db 101 GlyLeThrValSerIleSerGlyAlaGlyGluValIlePheThrPaspAsnSerValGly 120

QY 3886 TATTCCTCTTATCTACTGTGCCAACTCATCATCAACTCCGCTCTCT----- 3933  
 Db 121 TyrSerProLeuSerThrValProAlaSerThrProThrProProAlaProAla 140

QY 3934 -----CCAAAGTTAGTGTGCTCGGAAAGGGTCTATTTTCTGTA 3975  
 Db 141 AlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePheSerVal 160

QY 3976 GAGACTAGTTTGAGATCTCAGCGCTCAAAAAGGGTCATGTCGATTAATATCGCGG 4035  
 Db 161 GluThrSerLeuGluIleSerGlyValLysGlyValMetPheAspAsnAlaGly 180

QY 4036 AATTCGGACAGTTTTTCGAGGTAAAGTAATAATGCTGCTGTCGAGGAGCAGTGGG 4095  
 Db 181 AsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGlyGlySer-GI 200

QY 4096 TTCGGCTACACCATCAAGTACGACTTTTACAGTTAAAACTGTAAAGGGAAAGTTCTTT 4155  
 Db 200 ySerAlaThrProSer-----PheThrValLysAsnCysLysGlyLysValSerPh 218

QY 4156 CACAGATAACGTAGCTCTTCGGAGCGCGAGTGGTTTATAAGGCATTCGCTTTTCAA 4215  
 Db 218 eThrAspAsnValAlaSerCysGlyGlyValValValTyrLysGlyThrValLeuPheLys 238

QY 4216 AGCAATGAGGAGGCATATCTTCCGAGGGMAACACAGCATACGATCATTTAAGGATTCT 4275  
 Db 238 sAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLeuGlyIleLe 258

QY 4276 TGCTGCTACTAATCAGGATCAGAAATACGAGACAGGAGCGGTGAGAGTATTATTGCTC 4335  
 Db 258 uAlaAlaThrSerArgAspGlnAnThrGluThrGlyGlyGlyGlyValIleCysSe 278

QY 4336 TCCAGATGATTCTGTAAAGTTTGAAGCAATAAAGGTTCTATTGTTTGTGATTACAATT 4395  
 Db 278 rProAspAspSerValLysPheGluGlyAsnLysGlySerIleValPheAspTyrAsnPh 298

QY 4396 TGCAAAAGGACAGCGAGCGAGCATCCTTAACGAAAGAAATTC 4435  
 Db 298 eAlaLysGlyArgGlyGlySerIleLeuThrLysGluPhe 311

RESULT 11  
 Q83UW2\_CHLTR  
 ID Q83UW2\_CHLTR PRELIMINARY; PRT; 991 AA.  
 AC Q83UW2;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
 DE Polymorphic membrane protein H (Fragment).  
 GN Name=pmpH;  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OC NCBI\_TaxID=813;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22483673; PubMed=12595433;  
 RX DOI=10.1128/IAI.71.3.1200-1208.2003;  
 RA Stothard D.R., Toth G.A., Batteiger B.E.;  
 RT "Polymorphic membrane protein H has evolved in parallel with the three  
 disease-causing groups of Chlamydia trachomatis.";  
 RL Infect. Immun. 71:1200-1208(2003).  
 DR EMBL; AY184159; AAC030023.1; -; Genomic\_DNA.  
 DR EMBL; AY184163; AAC030027.1; -; Genomic\_DNA.  
 DR EMBL; AY184162; AAC030026.1; -; Genomic\_DNA.  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR InterPro; IPR005546; Auto transptbeta.  
 DR InterPro; IPR011427; ChlamPMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF07548; Chlam\_PMP; 6.  
 DR TIGRfam; TIGR01376; POMP repeat; 4.  
 DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
 FT NON\_TER 1  
 FT NON\_TER 991 991  
 SQ SEQUENCE 991 AA; 105133 MW; 09586A6137569398 CRC64;

Alignment Scores:  
 Pred. No.: 1,31e-87 Length: 991  
 Score: 1452.00 Matches: 286  
 Percent Similarity: 92.1% Conservative: 5  
 Best Local Similarity: 90.5% Mismatches: 10  
 Query Match: 18.4% Indels: 15  
 DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x Q83UW2\_CHLTR (1-991)

QY 3526 TCTCTCAAGTGTAAACGCTAATGTAAACACTCTCTTTAAGGAGACGATGTTACTTG 3585  
 Db 1 SerProGlnValLeuThrProAnValThrThrProPheLysGlyAspValTyrLeu 20

QY 3586 AATGGAGACTGCGCTTTTCTCAATGTCATGAGGAGCTGAAGAGTTTCGATTATCTCA 3645  
 Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleSer 40

QY 3646 GCTAATGGGACAAATTACGATTACCGGACAAACCATACATTATCATTTACAGATTCT 3705  
 Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHieThrLeuSerPheThrAspSer 60

QY 3706 CAAGGCCAGTCTCTCAAAATATGCTCTTCATTTACAGAGGAGACACTTACTCTGAGA 3765  
 Db 61 GlnGlyProValLeuGlnAnThrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80

QY 3766 GATTTTCGAGTCTGATGTCGAAAAATGTTCTTCGGGAGAAAAAGGAATGATCTCC 3825  
 Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100

QY 3826 GGGAAACGTCGATTTTCGGGAGCAGGAGTGTATTTCTCGGATACCTCCGNGGG 3885  
 Db 101 GlyLeThrValSerIleSerGlyAlaGlyGluValIlePheThrPaspAsnSerValGly 120

QY 3886 TATTCCTCTTATCTACTGTGCCAACTCATCATCAACTCCGCTCTCT----- 3933  
 Db 121 TyrSerProLeuSerThrValProAlaSerThrProThrProProAlaProAla 140

QY 3934 -----CCAAAGTTAGTGTGCTCGGAAAGGGTCTATTTTCTGTA 3975  
 Db 141 AlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePheSerVal 160

QY 3976 GAGACTAGTTTGAGATCTCAGCGCTCAAAAAGGGTCATGTCGATTAATATCGCGG 4035  
 Db 161 GluThrSerLeuGluIleSerGlyValLysGlyValMetPheAspAsnAlaGly 180

QY 4036 AATTCGGACAGTTTTTCGAGGTAAAGTAATAATGCTGCTGTCGAGGAGCAGTGGG 4095  
 Db 181 AsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGlyGlySer-GI 200

QY 4096 TTCGGCTACACCATCAAGTACGACTTTTACAGTTAAAACTGTAAAGGGAAAGTTCTTT 4155  
 Db 200 ySerAlaThrProSer-----PheThrValLysAsnCysLysGlyLysValSerPh 218

QY 4156 CACAGATAACGTAGCTCTTCGGAGCGCGAGTGGTTTATAAGGCATTCGCTTTTCAA 4215  
 Db 218 eThrAspAsnValAlaSerCysGlyGlyValValValTyrLysGlyThrValLeuPheLys 238

QY 4216 AGCAATGAGGAGGCATATCTTCCGAGGGMAACACAGCATACGATCATTTAAGGATTCT 4275  
 Db 238 sAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLeuGlyIleLe 258



```

Db      238 uphelyAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyAspAspLeuG 258
QY      4270 GATTCTTGCTGCTACTTAATCAGATCAGATACGAGACAGAGCGCGTGGAGGATTAT 4329
Db      258 yileLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyValI 278
QY      4330 TTGCTCTCCAGATGATCTGTTAAAGTTTGAAGCAATAAAGGTTCTATTGTTTTCATTA 4389
Db      278 eCySerProAspAspSerValysPheGluGlyAsnLysGlySerIleValPheAspTy 298
QY      4390 CAACTTTGCAAAAGGCGAGCGGAGCATCTTAACGAAAGAAATTC 4435
Db      298 rAsnPheAlaLysGlyArgGlySerIleLeuThrLysGluPhe 313

RESULT 13
Q84FV1_CHLTR
ID      Q84FV1_CHLTR PRELIMINARY; PRT; 991 AA.
AC      Q84FV1;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Polymorphic membrane protein H (fragment).
GN      Names:pmph;
OS      Chlamydia trachomatis.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=813;
[1]
NUCLEOTIDE SEQUENCE.
RX      DOI=10.1128/IAI.71.3.1200-1208.2003;
RA      Stothard D.R., Toth G.A., Basteiger B.E.;
RT      disease-causing groups of Chlamydia trachomatis.";
RL      Infect. Immun. 71:1200-1208(2003).
DR      EMBL; AY184156; AAO30024.1; -; Genomic DNA.
GO      GO:0019867; C:outer membrane; IEA.
DR      InterPro; IPR005546; Auto transportbeta.
DR      InterPro; IPR011427; ChlamPMP_M.
DR      InterPro; IPR003368; Chlamydia_PMP.
DR      InterPro; IPR002016; Peroxidase.
DR      Pfam; PF03797; Autotransporter; 1.
DR      Pfam; PF02415; ChlamPMP M; 1.
DR      Pfam; PF02415; Chlam_PMP; 6.
DR      TIGRFAMs; TIGR01376; POMP_repeat; 4.
DR      PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
FT      NON_TER 1
FT      NON_TER 991
SQ      SEQUENCE 991 AA; 105126 MW; 0004DC48387A2478 CRC64;

Alignment Scores:
Pred. No.: 1.31e-87 Length: 991
Score: 1452.00 Matches: 286
Percent Similarity: 92.1% Conservative: 5
Best Local Similarity: 90.5% Mismatches: 10
Query Match: 18.4% Indels: 15
DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x Q84FV1_CHLTR (1-991)

QY      3526 TCTCTCAAGTGTAAACGCTTAATGTAACCACTCCTTTTAAAGGAGACGATGTTTACTTG 3585
Db      1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValTyI 20
QY      3586 AATGAGACTCGCTTTTGTCAATGCTCTATCAGAGCTGAAGAGGTTTCGATTATCTCA 3645
Db      21 AsnGlyAspCysAlaPheValAsnValTyAlaGlyAlaGluAsnGlySerIleIleSer 40
QY      3646 GCTAATGCGCAATTTTAAACGATTACCGGACAAACCATATCATTTTACAGATTCT 3705
Db      41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60
QY      3706 CAAGGCCAGTCTTCAAAATATGCTTCAATTTTCAGCAGAGACACATTTACTCTGGA 3765

```

```

Db      61 GlnGlyProValLeuGlnAsnTyAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80
QY      3766 GATTTTTCGAGTCTCATGTTCTCGAAAAATGTTTCTTCGGAGAAAAAGGAATGATCTCC 3825
Db      81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100
QY      3826 GGGAAAAACGTCGATGATTTCGCGACGCGGAGGAGTATTTCTGGGATNACTCCGTCGGG 3885
Db      101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrAspAsnSerValGly 120
QY      3886 TATTTCTCTTATCTACTGTGCAACCTCATCATCACTCCGCTGCT----- 3933
Db      121 TySerProLeuSerIleValProAlaSerThrProThrProProAlaProAlaProAla 140
QY      3934 -----CCAACAGTTAGTGATGCTCGGAAAGGGTCTATTTT 3969
Db      141 ProAlaAlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePhe 160
QY      3970 TCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAAGGGTCAATGTTCCGATATAAT 4029
Db      161 SerValGluThrSerLeuGluIleSerGlyValLysLysGlyValMetPheAspAsnAsn 180
QY      4030 GCGGGAATTCGGAACAGTTCGAGGTAAAGATAATAATAATGCTGCTGCTGAGGCG 4089
Db      181 AlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsnAlaGlySerGlyGly 200
QY      4090 AGTGGGTCGCTACACCATCAAGTACGACTTTTACAGTTAAAAAAGCTGTAAGGGGAAAGT 4149
Db      201 Ser-GlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyLysVa 218
QY      4150 TTCTTTTCACAGATAACGTAGCTCTTTCGCGAGCGGAGTGGTTTATAAGGCAATGTCGT 4209
Db      218 IserPheThrAspAsnValAlaSerCysGlyGlyValValTyLysGlyThrValLe 238
QY      4210 TTTCAAAGCAATGAAGGAGGCAATATCTTCGAGGAGACACAGCATACATGATTTAAG 4269
Db      238 uphelyAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyAspAspLeuG 258
QY      4270 GATTCTTGCTGCTACTTAATCAGATCAGATCAGAGACAGAGCGCGTGGAGGAGTAT 4329
Db      258 yileLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyValI 278
QY      4330 TTGCTCTCCAGATGATCTGTTAAAGTTTGAAGCAATAAAGGTTCTATTGTTTTCATTA 4389
Db      278 eCySerProAspAspSerValysPheGluGlyAsnLysGlySerIleValPheAspTy 298
QY      4390 CAACTTTGCAAAAGGCGAGCGGAGCATCTTAACGAAAGAAATTC 4435
Db      298 rAsnPheAlaLysGlyArgGlySerIleLeuThrLysGluPhe 313

RESULT 14
Q83TU6_CHLTR
ID      Q83TU6_CHLTR PRELIMINARY; PRT; 993 AA.
AC      Q83TU6;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE      Polymorphic membrane protein H (fragment).
GN      Names:pmph;
OS      Chlamydia trachomatis.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=813;
[1]
NUCLEOTIDE SEQUENCE.
RX      DOI=10.1128/IAI.71.3.1200-1208.2003;
RA      Stothard D.R., Toth G.A., Basteiger B.E.;
RT      disease-causing groups of Chlamydia trachomatis.";
RL      Infect. Immun. 71:1200-1208(2003).
DR      EMBL; AY184156; AAO30020.1; -; Genomic DNA.
DR      EMBL; AY184158; AAO30022.1; -; Genomic DNA.
DR      EMBL; AY184157; AAO30021.1; -; Genomic DNA.

```

DR GO:0019867; C:outer membrane; IEA.  
DR InterPro; IPR005546; Auto\_transpbeta.  
DR InterPro; IPR011427; ChlamPMP\_M.  
DR InterPro; IPR003368; Chlamydia\_PMP.  
DR InterPro; IPR002016; Peroxidase.  
DR Pfam; PF03797; Autotransporter; 1.  
DR Pfam; PF07548; ChlamPMP\_M; 1.  
DR Pfam; PF02415; Chlam\_PMP; 6.  
DR TIGRFAMs; TIGR01376; POMP\_repeat; 4.  
DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 993 AA; 105683 MW; D20AABD706DEABA3 CRC64;  
Alignment Scores:  
Pred. No.: 1.88e-84 Length: 993  
Score: 1404.50 Matches: 279  
Percent Similarity: 90.9% Conservative: 9  
Best Local Similarity: 88.0% Mismatches: 13  
Query Match: 17.8% Indels: 16  
DB: 2 Gaps: 3  
US-10-701-844-1 (1-4435) x Q83TJ6\_CHLTR (1-993)  
QY 3526 TCTCTCAAGTGTAAACGCTAATGTAACCACTCTCTTTAAGGAGAGCATGTTACTTG 3585  
Db 1 SerProGlnValLeuThrProAsnValThrProPheLysGlyAspValTyrLeu 20  
QY 3586 AATGAGACTGCGCTTTTGTCAATGCTATGCGAGGCTGAAGAAGGTTCCATTATCTCA 3645  
Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleIleSer 40  
QY 3646 GCTAATGCGCAATTAACGATTACCGGACAAACCATACATTATCATGATTCT 3705  
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrHisSer 60  
QY 3706 CAAGGGCCAGTTCTCAAAATATGCTTCATTCAGCAGGAGACACTTACTCTGAGA 3765  
Db 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80  
QY 3766 GATTTTTCGAGTCTGATGTTCTCGAAAATGTTCTTCGGAGAAAAGGATGATCTCC 3825  
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100  
QY 3826 GGGAAAACGTAAGTATTTCCGGAGACGCGAAGTATTTTCGGGATAACTCCGTGGGG 3885  
Db 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheArgAspAsnSerValGly 120  
QY 3886 TATTCTCTTATCTACTGTGCGAACCTCATCATCACTCCGCTGCT----- 3933  
Db 121 TyrSerProLeuSerThrValProThrSerSerThrProAlaProAspProSer 140  
QY 3934 -----CCAAAGTGTAGTGTGCTCGAAAAGGGTCTATTTT 3969  
Db 141 ProThrAlaSerSerSerSerProThrValSerAspAlaArgGlnGlySerIlePhe 160  
QY 3970 TCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAGGGGTCAATGTTGATTAAT 4029  
Db 161 SerIleGluThrSerLeuGluIleSerGlyIleLysLysGluValMetPheAspAsn 180  
QY 4030 GCCGGGAATTCGGAACAGTTTTCGAGGTAAGATAATAATGCTGGTGGTGAGGC 4089  
Db 181 AlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGly 200  
QY 4090 AGTGGGTTCCGCTACACCATCAAGTACGCTTTTACAGTTAAAACTGTAAGGGAAAGT 4149  
Db 201 Ser-GlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyLys 218  
QY 4150 TTTCTTACAGATAACGTAAGCTCTTCGGAGGCGAGTGGTTTATAAGGATGTCGT 4209  
Db 218 LserPheThrAspAsnValAlaSerCysGlyGlyValValTyrLysGlyThrVal 238  
QY 4210 TTTCAAAGACATGAAGGAGCATATTCTTCGAGGGAACACAGCATACGATGATTAA 4269

Db 238 uphelyAspAsnGlnGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLeuG1 258  
QY 4270 GATTCTTGCTCTACTAATCAGGATCAG---AATACCGAGACAGAGCGGTGAGGAGT 4326  
Db 258 yileLeuAlaAlaThrSerGlnAspGlnSerAsnThrGluThrGlyGlyGlyAl 278  
QY 4327 TATTGCTCTCCAGATGATCTGTAAGTTTGAAGGCAATAAAGGTTCTATTGTTTGA 4386  
Db 278 aileCysSerProGluAspThrValThrPheGluGlyAsnLysGlySerIleValPheAs 298  
QY 4387 TTACAACCTTTGCAAAAGCGAGCGGAGCATCTTAACGAAGAATTC 4435  
Db 298 ptyrAsnPheAlaLysGlyArgGlyGlyIleLeuThrLysPhe 314  
RESULT 15  
Q84FV2\_CHLTR  
ID Q84FV2\_CHLTR PRELIMINARY; PRT; 993 AA.  
AC Q84FV2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Polymorphic membrane protein H (Fragment).  
GN Name=pmpH;  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP NEULBOTIDE SEQUENCE.  
RX MEDLINE=22483673; PubMed=12595433;  
RA DOI=10.1128/JAI.71.3.1200-1208.2003;  
RA Stothard D.R., Toth G.A., Batteiger B.E.;  
RT "Polymorphic membrane protein H has evolved in parallel with the three  
RT disease-causing groups of Chlamydia trachomatis.";  
RL Infect. Immun. 71:1200-1208(2003).  
DR EMBL; AY184155; AAC00019.1; -, Genomic\_DNA.  
DR GO:0019867; C:outer membrane; IEA.  
DR InterPro; IPR005546; Auto\_transpbeta.  
DR InterPro; IPR011427; ChlamPMP\_M.  
DR InterPro; IPR003368; Chlamydia\_PMP.  
DR InterPro; IPR002016; Peroxidase.  
DR Pfam; PF03797; Autotransporter; 1.  
DR Pfam; PF07548; ChlamPMP\_M; 1.  
DR Pfam; PF02415; Chlam\_PMP; 6.  
DR TIGRFAMs; TIGR01376; POMP\_repeat; 4.  
DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 993 AA; 105699 MW; 381340DD86400E39 CRC64;  
Alignment Scores:  
Pred. No.: 1.88e-84 Length: 993  
Score: 1404.50 Matches: 279  
Percent Similarity: 90.9% Conservative: 9  
Best Local Similarity: 88.0% Mismatches: 13  
Query Match: 17.8% Indels: 16  
DB: 2 Gaps: 3  
US-10-701-844-1 (1-4435) x Q84FV2\_CHLTR (1-993)  
QY 3526 TCTCTCAAGTGTAAACGCTAATGTAACCACTCTCTTTAAGGAGAGCATGTTACTTG 3585  
Db 1 SerProGlnValLeuThrProAsnValThrProPheLysGlyAspValTyrLeu 20  
QY 3586 AATGAGACTGCGCTTTTGTCAATGCTATGCGAGGCTGAAGAAGGTTCCATTATCTCA 3645  
Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleIleSer 40  
QY 3646 GCTAATGCGCAATTAACGATTACCGGACAAACCATACATTATCATGATTCT 3705  
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrHisSer 60  
QY 3706 CAAGGGCCAGTTCTTCAAAATATGCTTCATTCAGCAGGAGACACTTACTCTGAGA 3765





Score:	1365.50	Matches:	350
Percent Similarity:	49.9%	Conservative:	171
Best Local Similarity:	33.5%	Mismatches:	420
Query Match:	17.3%	Indels:	103
DB:	1	Gaps:	27

US-10-701-844-1 (1-4435) x PMP13\_CHLPN (1-973)

QY	382	ATGCAACGCTCTTCCATAAGTCTCTTCTTCAATGATCTTACGCTTATCTTCTGCTGCTCT	441
DB	1	MetLysThrSerLeuArgLysPheLeuIleSerThrLeuAlaProCysPheAlaSer	20
QY	442	TTAAATGGGGGGGATATGACAGCAGAAATCATGGTTCCTCAAGGAATTTACGATGGGGAG	501
DB	21	Thr-----AlaPheThrValGluValIleMetProSerGluAsnPheAspGlySer	37
QY	502	ACGTTAACTGTATCATCTCCCTATCTGTTATAGGAGATCCGAGTGGGACACTGTTT	561
DB	38	SerGlyLysIle---PheProTyThrThrLeuSerAspProArgGlyThrLeuCysIle	56
QY	562	TCGCGAGGAGGTTAACTAAAAATCTGTCAATCTTATGCAAGCTTTCGCTTTAAGT	621
DB	57	PheSerGlyAspLeuTyrlleAlaIleAlaLeuAspAsnAlaIleSerArgThrSerSer	76
QY	622	TGTTTGGGAACATTATAGGAGTCTTACTGTTTATAGGAGGAGGACACTCGTGA	681
DB	77	CysPheSerAsnArgAlaGlyAlaLeuGluIleLeuGlyLysGlyValPheSerPhe	96
QY	682	GAGAACATCGGACTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGA	738
DB	97	LeuAsnIleArgSerSerAlaAspGlyAlaAlaIleSerSerValIleThrGlnAsnPro	116
QY	739	-----CTGTTTACTATTGAGGCTTTTAAAGAAATTATCTTCCAAATTGCAATTCA	789
DB	117	GluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePheAspAsnCysGluSer	136
QY	790	TTACTTGGCGTACTGCTCTGCAACGACTAATAAGGGTAGCGACACTCCGACGACACA	849
DB	137	LeuThr-----SerAspThrSerAlaSerAsn	145
QY	850	TCATACCGTCTAATGGTACTATTATTCTTAAACAGATCTTTGTACTCAATAATGAG	909
DB	146	ValIleProHisAlaSerAlaIleTyAlaThrThrProMetLeuPheThrAsnAsnAsp	165
QY	910	AAGTCTCATATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGAGC	969
DB	166	SerIleLeuPheGlnTyrlenArgSerAlaGlyPheGlyAlaAlaIleArgGlyThrSer	185
QY	970	TTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAAAGAAATPACTGCTCAAGCTGAT	1029
DB	186	IleThrIleGluAsnThrLysLysSerLeuLeuPheAsnGlyAsnGlySerIleSerAsn	205
QY	1030	GGGGGAGCTGTCAAGTAGTACACAGTTCTCTGCTATGGCTTAACGAGGCTCTATTGGC	1089
DB	206	GlyGlyAlaLeuThrGlySerAlaAlaIleAsnLeuIleAsnAsnSerAlaProValIle	225
QY	1090	TTTGTAGCGAAATGTTTCAGAGTAAGAGGGGAGGATTCGTGTTTCAGGATGGGCAG	1149
DB	226	PheSerThrAsnAlaThrGlyIleTyrlleTyrlleTyrlleTyrlleTyrlle	241
QY	1150	CAGGGAGTGTATCATCTACTTCAACAGAAATCCAGTAGTAAGTTTTCAGAAATACT	1209
DB	242	---GlyGlySerMetLeuThrSer-----GlyAsnLeuSer	252
QY	1210	GGCGTAGTGTGATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC	1269
DB	253	GlyValLeuPheValAsnAsnSerSerArgSerGlyGlyAlaIleTyrlleAlaAsnGlyAsn	272
QY	1270	GTTCGCTTCTCGAATAATGGAACCTTGTTCTCAACAATGTTGCTTCTCTGTT---	1326
DB	273	ValThrPheSerAsnAsnSerAspLeuThrPheGlnAsnAsnThrAlaSerProGlnAsn	292
QY	1327	TACATTCGCTAAGCAACCAACAGGTGACAGGCTTCTTAATACAGTAATAATTACGGA	1386



```
Db 61 GlySerIleIleSerAlaAsnGlyGlnAsnLeuThrIleValGlyGlnAsnHisThrLeu 80
Qy 3691 TCATTACAGATTCTCAAGGCCAGTCTTCAAAATTAATGCTTCATTCACGAGGAG 3750
Db 81 SerPheThrAppSerGlnGlyProAlaLeuGlnAnCysAlaPheIleSerAlaGluGlu 100
Qy 3751 ACACCTACTCTGAGAGATTTTCGAGTCTGATGTTCTCGAAATATGTTCTTGGGAGAA 3810
Db 101 LysIleSerLeuArgAspPheSerSerLeuLeuPheSerLysAsnValSerCysGlyGlu 120
Qy 3811 AAGGAAATGATCTCCGGAAAAACCGTAGTATTTCGGGAGCAGCGGAGAGTATTTCTCG 3870
Db 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyLysAspSerIleValPheLys 140
Qy 3871 GATAACTCCGTTGGGTATTTCTCTTTACTTACTGTGCCAACCTCATCATCACTCCGCCT 3930
Db 141 AspAenSerValGlyTyrSerSerLeu-----ProSerValGlyGlnThrProThr 157
Qy 3931 CCTCAACAGTAGTAGTCTCGGAAAGGCTATTTTCTGTAGAGACTAGTTTGGAG 3990
Db 158 ThrProIleValGlyAspValLeuLysGlySerIlePheCysValGluThrGlyLeuGlu 177
Qy 3991 ATCTCAGCGCTCAAAAAGGGGTGATGTTTCATATATATGCGCGGAATTTTCGGAACAGTT 4050
Db 178 IleSerGlyValLysLysGluLeuValPheAspIleThrAlaGlyAsnPheGlyAlaVal 197
Qy 4051 TTTCAGGTGAGAATAATAATATGCTGTGTGGAGCAGTGGGTTCCGCTACACCCATC 4110
Db 198 PheCysSer-Arg-----AlaAlaGlnG1 205
Qy 4111 AAGTACGACTTTTACAGTTAAAACTGTAAGGGAAAGTCTTCTTCACAGTAAGTAGC 4170
Db 205 YAspThrThrPheThrValLysAspCysLysGlyLysIleLeuPheGlnAspAsnValG1 225
Qy 4171 CTCTTGGCGGCGGAGTGGTTTATAAGGATGTTGCTTTTCAAGACAATGAAGGAGG 4230
Db 225 YSerCysGlyGlyGlyValIleTyrLysGlyGluValLeuPheGlnAspAsnGluGlyG1 245
Qy 4231 CATATTCTCCGAGGAAACACAGCATACGATGATTATAAGGATTTCTGTCTACTATCA 4290
Db 245 uMetLeuPheArgGlyAsnSerAlaHisAspLeuGlyIleLeuAspAla---AsnPr 264
Qy 4291 GGATCAGATACGGAGACAGAGCGGTGGAGGAGTATTTGCTCTCAGATGATTTCTGT 4350
Db 264 oGlnProProThrGluValGlyGlyGlyValIleCysThrProGluLysThrVa 284
Qy 4351 AAAGTTTGAAGCAATAAAGTTCTATTGTTTGTGATTACAACTTTGCAGAAAGCAGAGG 4410
Db 284 lThrPheLysGlyAsnLysGlyProIleThrPheAspTyrAsnPheAlaLysGlyA-rg1 304
Qy 4411 CGAAGCATCTCAACGAAGAATTC 4435
Db 304 yGlyAlaIleGlnSerGlnThrPhe 312
RESULT 18
Q823X1_CHLCV
ID Q823X1_CHLCV PRELIMINARY; PRT; 942 AA.
AC Q823X1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family protein/autotransporter,
DE putative.
DE OrderedLocusNames=CCA00282;
GN Chlamydomophila caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83557;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
```

```
RA Heidelberg J.F., Holtzappple E.K., Khouri H.M., Federova N.B.,
RA Carty H.A., Umeyam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Heia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT *Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
RT Examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae.
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AF016995; AAP05033.1; -; Genomic_DNA.
DR TIGR; CCA00282; -.
GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto_transp_beta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
KW Complete proteome.
SQ SEQUENCE 942 AA; 100425 MW; E332BCB88507A912 CRC64;
Alignment Scores:
Pred. No.: 4,36e-68 Length: 942
Score: 1158.00 Matches: 326
Percent Similarity: 44.6% Conservative: 153
Best Local Similarity: 30.4% Mismatches: 402
Query Match: 14.7% Indels: 192
DB: 2 Gaps: 31
US-10-701-844-1 (1-4435) x Q823X1_CHLCV (1-942)
Qy 382 ATGCAACGCTTTCATTAAGTCTTCTTTCATTAAGTCTTAGCTTATCTTGTCTCTCT 441
Db 1 MetLysHisProValTyrTrpPheLeuValSerSerGlyLeuLeuAlaSer---ThrSer 19
Qy 442 TTAATCGGGGGGATATGCGACAGAAATCATGTTCTTCAAGAAATTTACGATGGGAG 501
Db 20 LeuSerPheAlaAlaThrValGlnGluThrLeuAsnSerSerAspSerTyrAsnGlyAsn 39
Qy 502 ACGTAACTCTGATCATTT-----CCCTATACTGTTTATA 534
Db 40 ThrAlaThrSerAlaPheGlnThrLysGluThrGlnAlaGlyAlaGluTyrThrCysGlu 59
Qy 535 GGAGATCCGAGTGGGACTACTGTTTTCGCGAGGAGTAAATTAATAAATCTTTGAC 594
Db 60 GlyAsn-----ValCysIleThrTyrAlaGlyLysGlyThr----- 71
Qy 595 AATTCATTGCGAGCTTTCCTTAAAGTGTGTTTGGGAACCTATTAGGAGTCTTACTGTT 654
Db 72 -----AlaLeuThrLysSerCysPheThrGluThrThrGluAsnLeuThrPhe 87
Qy 655 TTAGGGAGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACAAATGGGCGAGCT 714
Db 88 LeuGlyArgGlyTyrSerLeuCysPheAspAsnIleAsnThrThrAlaLysProAlaAla 107
Qy 715 CTAGTAATAGCGCTGCTGATGAGCTGTTTATTGAGGGTTTTAAAGAATTATCTCTT 774
Db 108 IleGluValSerAlaAlaAspLysThrLeuSerIleSerGlyPheSerLeuPheSerCys 127
Qy 775 TCCAAATTGCAATTCACTTACTGCGTACTGCTGCAACGACTAATAAGGTTAGCCAG 834
Db 128 SerAspCys-----ProGlyThrThrGlyGlnGly----- 138
Qy 835 ACTCCGACGACCAACATCTACACCGTCTAAATGGTACTATTATTCTAAACAGATCTTTTG 894
Db 139 -----AlaIleLysSerGlyGlyThrAlaThrPheGluAsnAspPheSer 153
Qy 895 TTACTCAATTAATGAGAAGTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGCT 954
Db 154 ValLeu-----PheLysLysAsnCysSerThrAlaAlaGlyGlyAla 167
Qy 955 ATAGATCCTAAGAGCTTAACGGTTCAAGGATTAAGCAAGCTTTGTGCTTCCAAAGAAAT 1014
```

Db 168 IleAsnCysLysGlyLeuThrLeuLysGlyThrSerGlyIleAlaAsnPheIleGluAsn 187  
Qy 1015 ACTGCTCAAGCTGATGAGGGGAGCTTGTCAAGTAGTACACAGTTTCTCGTATGCTAAC 1074  
Db 188 LysSerThrAspAsnGlyGlyAla 195  
Qy 1075 GAGGCTCTATTGCTTTGTAGCAATGTTGCAGGAGTAAGAGGGGAGGATTGCTGCT 1134  
Db 196 200  
Qy 1135 GTTCAGGATGGCAGCAGGAGTGCATCATCTACTTCAACAGAGATCCAGTAGTAAGT 1194  
Db 201 212  
Qy 1195 TTTTCCAGAAATCTGCGGTAGAGTTGTAGTGGAACTGAGCCGAGTAGGAGGAGATT 1254  
Db 213 PheSerGlyAsnThrSer 225  
Qy 1255 TACTCTACGGGAACTGTTGCTTCTGTAATATGGAATAACCTTGTCTCAACAATGTT 1314  
Db 226 HisSerAsnSerAlaValThrIleAlaAsnAsnHisArgValGluPhe 241  
Qy 1315 GCTTCTCTGTTTACATTGCTGCTAGCAACCAACAAAGTGGACAGGCTTCTAATACGAGT 1374  
Db 242 253  
Qy 1375 AATAATTACGAGATGAGGAGCTATCTTCTGTGAAGAATGGTGGCAAGCAGGATCCAAT 1434  
Db 254 262  
Qy 1435 AACTCTGGATCGATTCTCTTGTGAGGAGGAGTAGTCTTCTTTAGTACGATGAGCT 1494  
Db 263 AlaAlaProGluLeuLysPheGluGlyAsnAlaGlnLeuLeuPheLeuGluAsnSerSer 282  
Qy 1495 GCTGGGAAAGGGGAGCTATTATGCAAAAAGCTCTCGTGTCTAACTGTGGCCCTGTA 1554  
Db 283 GlnValSerGlyGlyAlaLeuTySerAsnLysLeuThrIleSerGlyGlyThrThr 302  
Qy 1555 CAAATTAAAGGAATATCGCTAATGAT 1602  
Db 303 ValPheAlaAsnAsnThrValThrAsnAlaAsnProMetGlyGlyAlaIleCysLeuAsp 322  
Qy 1603 ---GAATCTGGAGAGCTAGTTTATCTGCTGATATGAGATATTTTTCGATGGGAAT 1659  
Db 323 ThrThrSerGlyGlyCysSerLeuSerAlaAspLeuGlyAsnIlePheAspGlyAsn 342  
Qy 1660 CTTAAAGACACCAAGAGATGCTGCGGATGTTAATGCGCTAACTGTGCTCCACAA 1719  
Db 343 LysValIleThrSerGlyGlyAsn 356  
Qy 1720 GCCATTTTCGATGGATCGGAGGAGAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAG 1779  
Db 357 SerIleAspLeuAsnThrSerGlyLysPheThrGlnLeuArgAlaLysAspGlyPheGly 376  
Qy 1780 ATTCTCTTTAATGATCCATCGAGATGGCAAAACGGAATAACCGCAGCGCAGCTCTCC 1839  
Db 377 IleTyPheTyAspProIle---AlaAspAsnGlyAspAlaAsnAlaIle 392  
Qy 1840 AAATCTTAAATAATTAAGATGGTGAAGGA 1890  
Db 393 ---LeuAsnIleAsnAlaProGluAsnAlaThrThrTyAsnGlyArgValValPhe 410  
Qy 1891 GCT 1911  
Db 411 SerGlyGluThrLeuSerAlaThrGluLysThrGluAlaAspAsnLeuLysSerIlePhe 430  
Qy 1912 TACCAAAATGTTACGATAGACAAAGGAGGATGTTCTCTGTAAGAAAGGCAAAATATATCA 1971  
Db 431 LysGlnProValThrLeuSerAlaGlySerLeuIleLeuLysAspGlyValThrValGlu 450  
Qy 1972 GTGAATCTCTAAGTCACAGAGTGGGAGT---CTGTATATGGAAGCTGGGAGTACATGG 2028  
Db 451 AlaLysLysIleThrGlnThrAlaGlySerAlaValValMetAspAlaGlyThrThrLeu 470

Qy 2029 GATTTTGTAACTCCACCAACACCAACAGCCTCTCTGCCGCTAATATCATGATTCACGCTT 2088  
Db 471 482  
Qy 2089 TCCAATCTGCATTTGTCTCTTTCTTTGTTAGCAACAATGCAATGCAATCTCTCTCT 2148  
Db 483 ProAspLeuThrIleAsnValAlaSerPheGlyGlyGluGlyThr 498  
Qy 2149 ACCAATCTCCAGCGCAAGATTCTCATCTGCAGTCATTGCTAGCACAACATGCTGCT 2205  
Db 499 511  
Qy 2206 TCTGTTACAATTAGTGGCCTATCTTTTTCAGGATTTGGATGATPACAGCTTATGATAGG 2265  
Db 512 AlaLeuThrValThrAlaValSerPheIleAspAsp---AspGlyAsnGlyTyGlyTy 530  
Qy 2266 TATGATGGCTAGGTTCTAATCAAAAATCAATGCTCTCCGAAATTAACAGTTAGGGAATAG 2325  
Db 531 ProValPheSerLysThrArgAspPheAlaAspSerIleLeuLeuGluAlaAlaThrGly 550  
Qy 2326 CCCCAGCTAATGCCCATCA---GATTTGACTCTAGGAATGAGATGCTTAAGTATGCG 2382  
Db 551 ThrThrValThrAlaProAlaIleProThrThrProAspThrProSerAlaHisTyGly 570  
Qy 2383 TATCAAGGAAGCTGGAAGCTTGCCTGGGATCCATAACAGCAAAATAATGGTCTTATATCT 2442  
Db 571 TyGlnGlnAsnTrpThrIleAlaTrpAlaGlnGlyThrAlaGlyThrHisGluGlnLys 590  
Qy 2443 CTGAAGCTPACAGTAAACTGGGTATTAATCTCGGCTATTAATCTCGGCTGAGGAGTAGCTTTTG 2502  
Db 591 AlaThrLeuThrTrpThrGlyThrGlyTyIleProAsnProGluArgGlnAlaGlnLeu 610  
Qy 2503 GTTCCAATAGTTATGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATTC 2562  
Db 611 ValProAsnThrLeuTrpGlyAlaPheThrAspMetArgAlaLeuHisGlnLeuMetSer 630  
Qy 2563 GCAAGTGTGATGGCGCTCTTATTGTCGAGGATTTGGTCTTCTGAGATTTTCGAATTC 2622  
Db 631 ValSerAlaThrAspLeuGluGlnGluArgGlyLeuTrpGlyAlaAlaIleThrAspPhe 650  
Qy 2623 TTCTATCATGACCGGATGCTTTAGTCAGGATATCGGTATATATAGTGGGTTATTC 2682  
Db 651 LeuGlnArgLysLysThrSerLysLysTyArgHisValGlyValGlyTyAla 670  
Qy 2683 TTAGGAGCAACTCTCTACTTTGGATCA---TCGATGTTGCTCTAGCATTTACCGAAGTA 2739  
Db 671 ValGlyAlaSerValHisMetProThrGluAspLeuPheSerLeuAlaPheCysGlnPhe 690  
Qy 2740 TTTGTTAGATCTAAAGATTATGTTGTGTCGTTCCAAATCATCATCTGTCATAGGATCC 2799  
Db 691 PheAsnAsnAspLysAspPheValValSerLysAsnArgThrHisValTyAlaGlySer 710  
Qy 2800 GTTTAT---CTATCTACCCAAACAAAGCTTTATGTGATGCTTATTC 2841  
Db 711 LeuPhePheGluHisPheHisMetLeuHisProGlnAsnTyLysValGlySerLys 730  
Qy 2842 TTCGAGATGCGTTT 2862  
Db 731 PheProAlaPheLeuAlaAsnLeuProGluAsnValProMetIleLeuAsnIleLeu 750  
Qy 2863 GCTAGCTACGGTTCGGATCAGCATATGAAACCTCATATACA 2907  
Db 751 PheSerTySerHisAlaGluAsnAspMetLysThrArgTyThrLysArgTySerPro 770  
Qy 2908 ---TTTCAGAGGAGAGGAGTGTTCGTTGGGATAATACTGCTGGCTGGAGAG 2958  
Db 771 LysProValThrTyProGluValThrGlySerTrpGlyThrAsnCysPheGlyGlu 790  
Qy 2959 ATTGGAGCGGATTCAGGATTTGATCTCCATCTAAGCTCTATTTGAATGAGTTGCGT 3018  
Db 791 IleSerThrSerPheProIleGluLeuSerAspSerTyMetPhe---GluArgPheVal 809



1513 ATTTATGCCAAAGCTCTCGTTCGTAACCTGTCGCTGTACAAATTTTAAAGGAAT--- 1569  
1521 TTTCTGCTGATATGAGGATATTTATTTAGGAGAAATCTGGAGAGCTCAGT 1620  
1570 -----ATCGCTAAATGATCGTGGAGCGATTTATTTAGGAGAAATCTGGAGAGCTCAGT 1620  
312 ValThrHisAlaThrProLysGlyGlyAlaIleGlyIleAlaAlaSerGlyGlyCysSer 331  
1621 TTTATCTGCTGATATGAGGATATTTATTTAGGAGAAATCTGGAGAGCTCAGT 1680  
332 LeuThrAlaGluHisGlyAspIleThrPheAspAsnAsnLeuMetAlaThr---GlnAsp 350  
1681 AATGCTGCCGATTTAAATGCGCTAATCTGCTCACAAGCCATTTCCGATGGCATCGGGA 1740  
351 AsnAla-----ThrIleLysArgAsnAlaIleAsnIleGluGlyAsn 364  
1741 GCGAAATAAACGACATTTAAGAGCTAAAGCAGGCGCATCAGATTTCTTTAATATCCCATC 1800  
365 GlyLysPheValAsnLeuArgAlaAlaSerGlyLysThrIleSerPheTyrAspProIle 384  
1801 GAGATGGCAACCGAAATAAACGAGCGCGAGCTCTTCCAACTTCTTAAATAATTAACGAT 1860  
385 ThrVal---GluGlyAsn-----AlaAlaAspLeuLeuThrLeuAsnLys 398  
1861 GGTGAGGA-----TACACAGGGGATATTTTGTCT----- 1893  
399 AlaGluGlyAspLysThrTyrAsnGlyArgIlePheSerGlyGlyLysLeuThrGlu 418  
1894 -----AATGGAAGCAGTACTTTCTTACCAAAATGTTACGATPAGAG 1932  
419 GluGlnAlaAlaValAlaAspAsnLeuLysThrThrPheThrGlnProIleThrLeuAla 438  
1933 CAAGGAAGGATTTCTTCTGTAAGGCAAAATTAATCAGTGAATTTCTTAAGTCAGACA 1992  
439 AlaGlyGluLeuValLeuArgSerGlyValGluValGluAlaLysThrValGlnThr 458  
1993 GGTGGAGCTCG---TATATGGAAGCTGGGATACATGGGATTTCTTAACCTCCACAACCA 2049  
459 AlaGlySerLeuIleLeuMetAspAlaGlyThr----- 469  
2050 CCACAACAGCTCTCGCGCTAATCAGTTGATCAGCGCTTTTCCAATCTGCAATTTGTCTCTT 2109  
470 -----LysLeu 471  
2110 TCTTCTTTGTAGCAACAATGAGTACGATTCGAATCTCTTACCAATCTCTCAGCGCAGAT 2169  
472 SerAlaLysThrGluAspAlaThrLeuThrAsnLeuAlaIleAsnProAsnThrLeuAsp 491  
2170 TCTCATCTCT---GCAGTCATGTTAGCACAACTGCTGGT---TCTGTTACAATTAAGTGGG 2223  
492 GlyLysLysPheAlaValAlaAspAlaValAlaAlaGlyLysAsnValThrLeuSerGly 511  
2224 CCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTT 2283  
512 AlaIleGlyValIleAspProThrGlyLysPheTyrGlu----- 524  
2284 AATCAAAAATCAAT-----GTCTGGAATTTACAGTTA-----GGAGCT 2322  
525 AsnHisLysLeuAsnAspThrLeuAlaLeuGlyIleGlnLeuSerGlyLysGlySer 544  
2323 AAGCCCCAGCTAATGCCCCATCGATTTGACTCTAGGAAATGATGCTTAAGTATGGC 2382  
545 ValThrThrAsnValProSerHisVal---ValGlyValAlaGluThrHisTyrGly 563  
2383 TATCAAGGAGCTGAAGCTTGGCTGGATCTCTAATACAGCAATAATGCTCTTATATC 2442  
564 TyrGlnGlyAsnTrpSerValSerIlePheValLysAspAsn---AsnSerAspProLysThr 582  
2443 CTGAAAGCT-----ACATGAGCTAAACCTGGGTATTAATCTCGGCTGAGCGAGTACGT 2496  
583 GlnThrAlaIlePheThrTrpAsnLysThrGlyTyrValProAsnProGluArgAla 602  
2497 TCTTTGGTTCAAAATAGTTTATGGGATGCCATTTTAGATATACGATCTGGCATTCAGCA 2556

603 ProLeuValLeuAsnSerLeuTrpGlySerPheIleAspLeuArgSerIleGlnAspVal 622  
2557 ATTCAAGCAAGCTGGATGGG---CGCTCTATTGTCGAGGATTTATGCTTCTGAGTT 2613  
623 LeuGluArgSerValAspSerIleLeuGluThrArgArgGlyLeuTrpValSerGlyIle 642  
2614 TCGAATTTCTTCTATCATGACCGGATGCTTTAGTCTCAGGATATCGGTATATTAGTGGG 2673  
643 GlyAsnPhePheHisLysAspArgAsnAlaGluAsnArgLysPheArgHisIleSer 662  
2674 GGTATTTCTTCTAGGAGCAACTCTCTACTTTTGA---TCATCGATTTTGGTCTACATTT 2730  
663 GlyTyrValLeuGlyAlaThrThrAsnThrSerArgGluAspSerLeuSerValAlaPhe 682  
2731 ACCGAGCTATTGTTAGTATCTAAAGATTTAGTGTGTCGTTCCAAATCATCATCTGCTGC 2790  
683 CysGlnLeuPheAlaLysAspLysAspTyrLeuValSerLysAsnAlaAlaAsnValTyr 702  
2791 ATAGGATCCGTTTAT-----CTATCT 2811  
703 AlaGlySerValTyrTyrGlnHisValSerLysPheAspAspLeuThrArgLeuPheAsn 722  
2812 ACCCAACAAGCTTTATGTGGATCTCTAT-----TTGTTCCGAGATCGGTTT 2856  
723 GlyProAsnThrCysCysSerGlyPheSerLysGluIleProIlePheLeuAspAlaGln 742  
2857 ATCCGTGTAGTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGTCAGAG 2916  
743 Ile-----ThrTyrCysHisThrAlaAsnAsnMetThrThrSerTyrThrAspTyrPro 760  
2917 GAGAGCGATGTTGCTGGGATATAATCTCTCTGGCTGGAGAGATTTGGAGCGGATTTACC 2976  
761 GluValLysGlySerTrpGlyAsnAspThrLeuGlyLeuThrLeuSerThrSerValPro 780  
2977 ATTGATATTCTCAATGCTCTTATTGTAAGTATGCTGCTCTTCTGTCAGCTGAG 3036  
781 IleProValPheSerSerSerIlePhe---AspSerTyrAlaProPheAlaLysLeuGln 799  
3037 TTTTCTTATGCGCATCATCAATCTTTTACAGAGGAGCGATCAAGCTCGGCATTTCAAG 3096  
800 ValValTyrAlaHisGlnAspAspPheLysGluProThrThrGluGlyArgValPheGlu 819  
3097 AGCGGACATCTCTAAATCTATCAGTTCTGTTGGAGTGAAGTTTGCATGCTTCTAGT 3156  
820 SerSerAspLeuAsnValSerValProIleGlyLysPheGlyLysLeuSerTyr 839  
3157 ACACATCTCTAATAATAGCTTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATC 3216  
840 GlyLysArgSerAlaTyrAspLeuThrLeuMetTyrIleProAspValTyrArgHisAsn 859  
3217 TCTGTTACTGAGCAACGCTCTCTATCCCATCAAGACACATGACACAGATGCTTTCAT 3276  
860 ProSerCysMetThrGlyLeuAlaIleAsnAspValSerTrpLeuThrThrAlaThrAsn 879  
3277 TTAGCAAGACATCGAGTGTGTTAGGATCTATCTATGCTTCTCTTAAACAAGTATATA 3336  
880 LeuAlaArgGlnAlaPheIleValArgAlaGlyAsnHisIleAlaLeuThrSerGlyVal 899  
3337 GAAGTATATGGCCATCGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGGT 3396  
900 GluMetPheSerGlnPheGlyPheGluLeuAspSerSerSerArgAsnTyrAsnValAsp 919  
3397 GCAGGAAGTATAGTCCGGTTC 3417  
920 LeuGlyAlaLysValAlaPhe 926  
RESULT 20  
QSL6J3\_CHLAB PRELIMINARY; PRT; 926 AA.  
ID QSL6J3\_CHLAB  
AC QSL6J3;  
DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)





QY 2032 TTTGTAACCTCCACCAACCACACAGCCTCTCGCGCTTAATCACTGATCAGCGCTTCC 2091  
DB 469 ----- 469  
QY 2092 AATCTGCATTTGCTCTCTTCTTTCTTTGTAGCAAAATGCAATGACGATCTCCTACC 2151  
DB 470 -----LysLeuSerAlaLysThrGluAspAlaThrLeuThrAsnLeuAlaIle 485  
QY 2152 AATCTCCAGCCGCAAGATCTCATCTCT---GCAGTCATTGGTAGACCAACTCTCGT--- 2205  
DB 486 AsnProAsnThrLeuAspGlyLysLysPheAlaValValAspAlaValAlaAlaGlyLys 505  
QY 2206 TCTGTTACAAATTAGTGGGCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGG 2265  
DB 506 AsnValThrLeuSerGlyAlaIleGlyValIleAspProThrGlyLysPheTyrGlu--- 524  
QY 2266 TATGATTGGCTAGGTTCTAATCAAAAATCAAT-----GTCTGAAATATACAG 2313  
DB 525 -----AsnHisLysLeuAsnAspThrLeuAlaLeuGlyGlyIleGln 538  
QY 2314 TTA-----GGGACTAAGCCCCAGCTAATGCCCATCATGATTTGACTTAGGGAAT 2364  
DB 539 LeuSerGlyLysGlySerValThrThrAsnValProSerHisVal---ValGlyVal 557  
QY 2365 GAGATGCCATAAGTATGCTATCAGGAAGCTGGAAGCTTGGCTGGATCCTAATACAGCA 2424  
DB 558 AlaGluThrHisTyrGlyTyrGlnGlyAsnTyrSerValSerTyrValLysAspAsn--- 576  
QY 2425 AATAATGGTCTTATCTCTGAAAGCT-----ACATGACTAAAACTGGGTATAATCCT 2478  
DB 577 AsnSerAspProLysThrGlnThrAlaIlePheThrTyrAsnLysThrGlyTyrValPro 596  
QY 2479 GGGCTTCAGCGAGTAGTCTCTTTGGTTCCAAATAGTTTATGGGATPCCATTTAGATATA 2538  
DB 597 AsnProGluArgAlaProLeuValLeuAsnSerLeuTyrGlySerPheIleAspLeu 616  
QY 2539 CGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGG---CGCTCTTATTTGTCAGCA 2595  
DB 617 ArgSerIleGlnAspValLeuGluArgSerValAspSerIleLeuGluThrArgArgGly 636  
QY 2596 TTATGGGTTCTCGAGTTCTGAAATTTCTCTATCATGACCGCATGCTTTAGTCAAGCA 2655  
DB 637 LeuTyrValSerGlyIleGlyAsnPheHisLysAspArgAsnAlaGluAsnArgLys 656  
QY 2656 TATCGGTATATTAGTGGGGTTATCTCTTAGGAGCAAACTCTCTACTTTGCA---TCATCG 2712  
DB 657 PheArgHisIleSerSerGlyTyrValLeuGlyAlaThrThrAsnThrSerArgGluAsp 676  
QY 2713 ATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCTTAAAGATTATGATGTCGT 2772  
DB 677 SerLeuSerValAlaPheCysGlnLeuPheAlaLysAspLysAspTyrLeuValSerLys 696  
QY 2773 TCCATCATCATCTTGCATAGGATCGTTAT----- 2805  
DB 697 AsnAlaAlaAsnValTyrAlaGlySerValTyrTyrGlnHisValSerLysPheAspAsp 716  
QY 2806 -----CTATCTACCCCAACAGCTTTATGTGGATCCTAT----- 2838  
DB 717 LeuThrArgLeuPheAsnGlyProAsnThrCysCysSerGlyPheSerLysGluIlePro 736  
QY 2839 TTGTTCCGAGATCGGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898  
DB 737 IlePheLeuAspAlaGlnIle-----ThrTyrCysHisThrAlaAsnAsnMetThrThr 754  
QY 2899 TCATATACATTTCAGAGAGGAGCGATGTTCTGTTGGATATAACTCTCGGCTGGAGAG 2958  
DB 755 SerTyrThrAspTyrProGluValLysGlySerTyrGlyAsnAspThrLeuGlyLeuThr 774  
QY 2959 ATTGGAGCGGATTTACCGATTGTGATTACTTCCATCTAAGTCTTATTTGAATGATGTCGT 3018  
DB 775 LeuSerThrSerValProIleProValPheSerSerSerIlePhe---AspSerTyrAla 793

QY 3019 CCTTTCTGCAAGCTGAGTTTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGCGCAT 3078  
DB 794 ProPheAlaLysLeuGlnValTyrAlaHisGlnAspAspPheLysGluProThrThr 813  
QY 3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTTAATCTATCATGTTCTCTGTTGGAGTGAAG 3138  
DB 814 GluGlyArgValPheGluSerSerAspLeuLeuAsnValSerValProIleGlyIleLys 833  
QY 3139 TTTGATCGATGTTCTAGTACACATCTTAATAATATAGCTTTTATGGCGGCTTATATGCT 3198  
DB 834 PheGluLysLeuSerTyrGlyGluArgSerAlaTyrAspLeuThrLeuMetTyrIlePro 853  
QY 3199 GATGCTTATCGCACCATCTCTGTTACTGAGACAGCTCTCTATCCATCAGAGACATGG 3258  
DB 854 AspValTyrArgHisAsnProSerCysMetThrGlyLeuAlaIleAsnAspValSerTyr 873  
QY 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGGAGTTGTTGGTAGAGGATCTATGATGCT 3318  
DB 874 LeuThrThrAlaThrAsnLeuAlaArgGlnAlaPheIleValArgAlaGlyAsnHisIle 893  
QY 3319 TCTCTAAACAAGTAATATAAGATATATGCGCATGGAAGATATGATATCGAGATGCTTCT 3378  
DB 894 AlaLeuThrSerGlyValGluMetPheSerGlnPheGlyPheGluLeuArgSerSerSer 913  
QY 3379 CGAGGCTATGTTGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417  
DB 914 ArgAsnTyrAsnValAspLeuGlyAlaLysValAlaPhe 926  
RESULT 21  
PMP10 CHLPN STANDARD; PRT; 928 AA.  
AC Q9RB65; O86163; Q9RB64; Q9S6P2;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable outer membrane protein pmp10 precursor (Polymorphic membrane  
protein 10) (Outer membrane protein 5).  
GN Name=pmp10; Synonyms=omp5; OrderedLocusNames=CP0303, CpB0467;  
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.  
OX NCBI\_TaxID=833558;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CWL029 / VR1310;  
RX MEDLINE=20007584; PubMed=10539856;  
RA Christensen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,  
Madisen A.S., Knudsen K., Falk E., Birkelund S.;  
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their  
role in immunopathogenicity."  
RL Am. Heart J. 138:S491-S495(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=AR39;  
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,  
Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,  
Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,  
McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
pneumoniae AR39."  
RL Nucleic Acids Res. 28:1397-1406(2000).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=J138;  
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;  
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
from Japan and CWL029 from USA."  
RL Nucleic Acids Res. 28:2311-2314(2000).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN-TW-183;  
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensach K.W., Schaefer K.P.,  
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;  
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with  
RT other Chlamydia strains based on whole genome sequence analysis.";   
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RN (5)  
RP NUCLEOTIDE SEQUENCE OF 1-914.  
RC STRAIN-CW029/VR-1310;  
RX MBLINS-99081766; PubMed:9864239;  
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;  
RT "Identification of two novel genes encoding 97- to 99-kilodalton outer  
RT membrane proteins of Chlamydia pneumoniae.";   
RL Infect. Immun. 67:375-383(1999).  
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)  
CC (Potential).  
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC EMBL: AJ133034; CAB37071.1; -; Genomic DNA.  
DR EMBL: AB020192; AAF38160.1; -; Genomic DNA.  
DR EMBL: BA000008; BAA98657.1; -; Genomic DNA.  
DR EMBL: AB017158; AAF98398.1; -; Genomic DNA.  
DR EMBL: AJ001311; CAA04671.1; -; Genomic DNA.  
DR PIR: G81591; G81591.  
DR PIR: G86546; G86546.  
DR PHCI-2DPAGE; O86163; -.  
DR TIGR: CP0303; -.  
DR InterPro: IPR005546; Auto transportbeta.  
DR InterPro: IPR006315; Autotransporter.  
DR InterPro: IPR011427; ChlamPMP M.  
DR InterPro: IPR003368; Chlamydia\_PMP.  
DR InterPro: IPR006826; PPH1.  
DR Pfam: PF03797; Autotransporter; 1.  
DR Pfam: PF02415; Chlam PMP; 7.  
DR Pfam: PF07548; ChlamPMP M; 1.  
DR SMART; SM00710; PPH1; 4.  
DR TIGRPfam; TIGR01414; autotrans barl; 1.  
DR TIGRPfam; TIGR01376; POMP\_repeat; 5.  
KW Membrane; Multigene family; Outer membrane; Signal.  
FT SIGNAL 1 25 Potential.  
FT CHAIN 26 928 Probable outer membrane protein pmp10.  
FT CONFLICT 305 305 Missing (in Ref. 4).  
SQ SEQUENCE 928 AA; 97230 MW; 0590D5206A1DD0E1 CRC64;  
Alignment Scores:  
Pred. No.: 2,65e-64 Length: 928  
Score: 1101.00 Matches: 324  
Percent Similarity: 44.5% Conservative: 150  
Best Local Similarity: 30.5% Mismatches: 402  
Query Match: 14.0% Indels: 188  
DB: 1 Gaps: 34  
US-10-701-844-1 (1-4435) x PMP10\_CHLNP (1-928)  
QY 382 ATGCAACGCTTTCCATAAGTCTTTCTTCAATGATTCTAGCT--TATTCCTGCTGC 438  
Db |||||:|||||  
QY 1 MetLysSerGlnPheSerTrpLeuValLeuSerSerThrLeuAlaCysPheThrSerCys 20  
Db |||||:|||||  
QY 439 TCTTTAAATGGGGGGGATATGCAGCAAAATCATGGTCTCTCAAGGAATTTACAGGG 498  
Db |||||:|||||  
QY 21 SerThrValPheAlaAla---ThrAlaGluAsnIleGlyProSerAspSerPheAspGly 39  
Db |||||:|||||  
QY 499 GAGACGTTAACTGTATCATTTCCCTATATCTGTATATAGAGATCCGAGTGGACTACTGTT 558  
Db |||||:|||||  
QY 40 SerThrAsnThrGlyThr-----TyrThrProLysAsnThrThrThrGlyIleAspTyr 57  
Db |||||:|||||  
QY 559 TTTTCGCGAGGAGTAAACATTAAATAATCTTGACATTTCTATTCGAGCTTTGCGTTTA 618  
Db |||||:|||||

Db |||||:|||||  
58 ThrLeuThrGlyAspIleThrLeuGlnAsnLeuGlyAspSer---AlaAlaLeuThrLys 76  
QY |||||:|||||  
619 AGTTGTTTGGGAACCTATTATTAGGAGGTTTACTGTTTATGGGAGAGGACACTCGTGTACT 678  
Db |||||:|||||  
77 GlyCysPheSerAspThrThrGluSerLeuSerPheAlaGlyLysGlyTyrSerLeuSer 96  
QY |||||:|||||  
679 TTCGAGAACATACGGACTTCTCAAAATGGGGGAGCTTAAAGTAAATAGCGGCTGCTGATGGA 738  
Db |||||:|||||  
97 PheLeuAsnIleLysSerSerAlaGluGlyAlaAlaLeuSer---ValThrThrAspLys 115  
QY |||||:|||||  
739 CTGTTTACTATGAGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTTACTTACTGCC 798  
Db |||||:|||||  
116 AsnLeuSerLeuThrGlyPheSerSerLeuThrPhe-----LeuAlaAla 130  
QY |||||:|||||  
799 GTACTGCTCTGCACAGCACTAAATAGGGTAGCCAGACTCCGACGACAAACATCTACACCG 858  
Db |||||:|||||  
131 ProSerValIleThrThrProSerGly----- 140  
QY |||||:|||||  
859 TCTAATGCTACTATTATTCTTAAACAGATCTTTTGTGTACTCAATATGAGAATTCCTCA 918  
Db |||||:|||||  
141 ---LysGlyAlaValLysCysGlyGlyAspLeuThrPheAspAsnAsnGlyThrIleLeu 159  
QY |||||:|||||  
919 TTCTATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACCGTT 978  
Db |||||:|||||  
160 PheLysGlnAspTyrCysGluGluAsnGlyGlyAlaIleSerThrLysAsnLeuSerLeu 179  
QY |||||:|||||  
979 CAAGGAATTAGCAAGCTTTGTGTTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCT 1038  
Db |||||:|||||  
180 LysAsnSerThrGlySerIleSerPheGluGlyAsnLysSerSerAla----- 195  
QY |||||:|||||  
1039 TGTCAGTAGTCACCAGTTCTCTGCTATGCTTAACGAGCTCTATTCGCTTTGTAGCG 1098  
Db |||||:|||||  
195 ----- 195  
QY |||||:|||||  
1099 AATGTTGCAGGAGTAAGAGGGGAGGATTGCTGCTGTTTCAGGATGGGACAGGAGGTG 1158  
Db |||||:|||||  
196 -----ThrGlyLysLysGlyGlyAlaIleCysAla----- 205  
QY |||||:|||||  
1159 TCATCATCTACTTCAACAGCAAGATCCAGTAGTAGTAAAGTTTTTCCAGAAATATCTGCG--GTA 1215  
Db |||||:|||||  
206 -----ThrGlyThr-----ValAspIleThrAsnAsnThrAlaProThr 218  
QY |||||:|||||  
1216 GAGTTTCATGGAACTAGTACCCGA---GTAGGAGGAGGATTTACTCTCTAGGGGAAAGCTT 1272  
Db |||||:|||||  
219 LeuPheSerAsnAsnIleAlaGluAlaAlaGlyGlyAlaIleAsnSerThrGlyAsnCys 238  
QY |||||:|||||  
1273 GCTTTCCTGAATTAATGAAAACCTGTTCTCAACAAATGTTGCTTCTCCTGTTTACATT 1332  
Db |||||:|||||  
239 ThrIleThrGlyAsnThrSerLeuValPheSerGluAsn----- 251  
QY |||||:|||||  
1333 GCTGCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAGTAAATATTACGAGATGGA 1392  
Db |||||:|||||  
252 -----SerValThrAlaThrAlaGlyAsnGly 260  
QY |||||:|||||  
1393 GGAGCTATCTTCTGTAGAAATGGTGGCAAGCAGGATCCAACTCTGGATCATGTTTCC 1452  
Db |||||:|||||  
261 GlyAlaLeu-----SerGlyAspAlaAspValThr 270  
QY |||||:|||||  
1453 TTTGATGAGGAGGAGTAGTTTCTTTTAGTAGCAATGCTAGCTGCTGGGAAAGGGGAGCT 1512  
Db |||||:|||||  
271 IleSerGlyAsnGlnSerValThrPheSerGlyAsnGlnAlaValAlaAsnGlyGlyAla 290  
QY |||||:|||||  
1513 ATTTATGCCAAAAGCTCTCGGTTGCTTAAC-----TGTGGCCCTGTACAAATTTTAAAG 1566  
Db |||||:|||||  
291 IleTyrAlaLysLysLeuThrLeuAlaSerGlyGlyGlyGlyIleSerPheSerAsn 310  
QY |||||:|||||  
1567 AATATC-----GCTAATGATGGTGGAGCGGATTATTATTAGGGAATCTGGA 1611  
Db |||||:|||||  
311 AsnIleValGlnGlyThrThrAlaGlyAsnGlyAlaIleSerIleLeuAlaAlaGly 330  
QY |||||:|||||  
1612 GAGCTCATGTTTATCTGCTGATTTATGGAGATATTATTTTCGATGGGAAT----- 1659  
Db |||||:|||||

Db 331 GluCySerLeuSerAlaGluAlaGlyAspIleThrPheAsnGlnGlyAlaValAla 350  
 QY 1660 ----CTTAAAGAACCCAAAGAGATGTCGGGATGTTAATGGCGTAACTGTGCTCA 1716  
 Db 351 ThrThrProGlnThrThrLysArgAsnSerIleAspIle----- 363  
 QY 1717 CAAGCCATTTCGATGGATCGGAGGCGAAATAACACATTAAGACCTAAAGCAGGCGCAT 1776  
 Db 364 -----GlySerThrAlaLysIleThrAsnLeuArgAlaIleSerGlyHis 378  
 QY 1777 CAGATTCTCTTAATGATCCCATCGAGATGGCAACGGAATAACACGAGCGCATGCT 1836  
 Db 379 SerIlePhePheThrAspProIleThrAla-----AsnThrAlaAlaAspSer 394  
 QY 1837 TCCAAACTCTAAATAATTACGATGCTGAAGG-----TACACAGGGGAT 1881  
 Db 395 ThrAspThrLeuAsnLeuAsnLysAlaAspAlaGlyAsnSerThrAspThrAspThrGlySer 414  
 QY 1882 ATGTGTTTTGCT-----AATGGAAGC 1902  
 Db 415 IleValPheSerGlyGluLysLeuSerGluAspGluAlaLysValAlaAspAsnLeuThr 434  
 QY 1903 AGTACTTTGACCAAAATGTTACGATAGACCAAGGAGATTGCTCTCGTGAAGGCA 1962  
 Db 435 SerThrLeuLysGlnProValThrLeuThrAlaGlyAsnLeuValLeuLysArgGlyVal 454  
 QY 1963 AATTATCAGTGAATCTCTAAGTCACACA--GGTGGAGTCTGTATATGGAAGCTGG 2019  
 Db 455 ThrLeuAspThrLysGlyPheThrGlnThrAlaGlySerSerValIleMetAspAlaGly 474  
 QY 2020 AGTACATGGGATTTGTTAACTCCACACCAACACAGCGCTCTCGCGCTAATCAGTTG 2079  
 Db 475 ThrThrLeu-----LysAlaSerThrGluGlu 483  
 QY 2080 ATCAGCTTTCCAAATGCTGCTCTCTTTCTTTGTTAGCAACAATGCGTTACG 2139  
 Db 484 ValThrLeuThrGlyLeuSerIleProValAspSerLeu----- 496  
 QY 2140 AATCCTCTACCAATCTCCAGCGCAAGATTCATCCTCGATCATCTGTGACACAACT 2199  
 Db 497 -----GlyGluGlyLysValValIleAlaAlaSerAlaAla 509  
 QY 2200 GCTGGTCTGTACAAATGATGGCGCTATCTTTTTCGAGGATTTGGATGATACAGCTTAT 2259  
 Db 510 SerLysValAlaLeuSerGlyProIleLeuLeuLeuAspAsnGlnGlyAsnAlaThr 529  
 QY 2260 GATAGGTATGATGGCTAGTCTTAATCAAAAATCAATGCTCTGAAATACAG---TTA 2316  
 Db 530 GluAsnHisAsp---LeuGlyLysThrGlnAspPheSerPheValGlnLeuSerAlaLeu 548  
 QY 2317 GGGACTAAGCCCCCAGCTAATGCCCATCAGATTGATCTAGGAATGAGATGCCTAAG 2376  
 Db 549 GlyThrAlaThrThrAspValProAlaValProThrValAlaThrPro---ThrHis 567  
 QY 2377 TATGGCTATCAAGCAAGCTGGAAGCTGCGGGATCCTAATACAGCAAAATAATGTCCT 2436  
 Db 568 TyrGlyThrGlnGlyThrThrPglyMetThrThrValAspAspThrAlaSerThr---Pro 586  
 QY 2437 TATACTCTGAAGCTACA-----TGGACTAAAATCGGTATATCTCGGCGCTGAGCGA 2490  
 Db 587 LysThrLysThrAlaThrLeuAlaThrThrAsnThrGlyThrLeuProAsnProGluArg 606  
 QY 2491 GTAGCTCTTTGGTTCCAAATGATTTATGGGATCCATTTTATAGATATACGATCTGCCAT 2550  
 Db 607 GlnGlyProLeuValProAsnSerLeuThrPglySerPheSerAspIleGlnAlaIleGln 626  
 QY 2551 TCAGCAATCAAGCAAGTGTGGTGGGCTCTTATTGTCGAGGATATTGGTTCTGGA 2610  
 Db 627 GlyValIleGluArgSerAlaLeuThrLeuCySerAspArgGlyPheThrAlaAlaGly 646  
 QY 2611 GTTTCGAATTTCTTATCATGACCGCGATGCTTTAGTCTAGGATATCGGTATATTAGT 2670  
 Db 647 ValAlaAsnPheLeuAspLysAspLysLysGlyGluLysArgLysThrArgHisLysSer 666

QY 2671 GGGGTTTATTCCTTA---GGAGCAAACTCTACTTTGGATCATCGATGTTGGTCTAGCA 2727  
 Db 667 GlyGlyThrAlaIleGlyAlaAlaGlnThrCysSerGluAsnLeuIleSerPheAla 686  
 QY 2728 TTTACCGAAGTATTTGGTAGATCTAAAGATATATGATGTGTCGTTCCTCAATCATGCT 2787  
 Db 687 PheCysGlnLeuPheGlySerAspLysAspPheLeuValAlaLysAsnHisThrAspThr 706  
 QY 2788 TCATAGGATCCGTTTATCTATCT-----ACC 2814  
 Db 707 TyrAlaGlyAlaPheThrIleGlnHisIleThrGluCysSerGlyPheIleGlyCysLeu 726  
 QY 2815 CAACAAGCTTTATGTGATCTCTATTGTTGGAGATGCGTTTATC-----CGTCTAGC 2868  
 Db 727 LeuAspLysLeuProGlySerTrpSerHisLysProLeuValLeuGluGlyGlnLeuAla 746  
 QY 2869 TACGGGTTTGGGAATCAGCATATGAACCTCATATACATTTTCAGAGGAGCGCATGTT 2928  
 Db 747 TyrSerHisValSerAsnAspLeuLysThrLysThrAlaThrProGluValLysGly 766  
 QY 2929 CGTTGGATAATAACTGCTCTGCGAGAGATTGGAGCGGATACCGATTGTGATTACT 2988  
 Db 767 SerTrpGlyAsnAsnAlaPheAsnMetMetLeuGlyAlaSerSerHisSerTyr----- 784  
 QY 2989 CCATCTAAGCTCTAT---TTGAATGAGTTGCGTCTCTTCGTGCAAGCTGAGTTTCTTAT 3045  
 Db 785 ProGluThrLeuHisCysPheAspThrTyrAlaProTyrIleLysLeuAsnLeuThrTyr 804  
 QY 3046 GCGGATCATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCGATTCAGAGCGGACAT 3105  
 Db 805 IleArgGlnAspSerPheSerGluLysGlyThrGluGlyArgSerPheAspAspSerAsn 824  
 QY 3106 CTCCTAAATCATGCTCTCTGAGGAGGAGTTTGTGATGCGATGTTCTAGTACACATCCT 3165  
 Db 825 LeuPheAsnLeuSerLeuProIleGlyValLysPheGluLysPheSerAspCysAsnAsp 844  
 QY 3166 AATAAATATAGCTTTATGCGGCTTATATCTGTGATGCTTATGCGCACCATCTCTGCTACT 3225  
 Db 845 PheSerTyrAspLeuThrLeuSerTyrValProAspLeuIleArgAsnAspProLysCys 864  
 QY 3226 GAGACAACGCTCTATCCCATCAAGAGACATGACAGATGCGCTTCTTATTAGCAAGA 3285  
 Db 865 ThrThrAlaLeuValIleSerGlyAlaSerTrpGluThrTyrAlaAsnAsnLeuAlaArg 884  
 QY 3286 CATGGATGTTGTTAGAGGATCTATGATGCTTCTCTAACAGTAAATATAGAGTATAT 3345  
 Db 885 GlnAlaLeuGlnValArgAlaGlySerHisTyrAlaPheSerProMetPheGluValLeu 904  
 QY 3346 GGCCATGGAAGATATCAGTATCGAGATGCTTCTCGAGGCTATGTTGATGTCAGGAAGT 3405  
 Db 905 GlyGlnPheValPheGluValArgGlySerSerArgIleTyrAsnValAspLeuGlyGly 924  
 QY 3406 AGAGTCCGGTTC 3417  
 Db 925 LysPheGlnPhe 928

RESULT 22

Q823W9 CHLCV  
 ID Q823W9 CHLCV PRELIMINARY; PRT; 926 AA.  
 AC Q823W9;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polymorphic outer membrane protein G family  
 DE protein/autotransporter.  
 DE OrderedLocusNames=CCA00284;  
 GN Chlamydophila caviae.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.  
 RN NCBI\_Taxid=83557;  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=GPIC;

```

RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapple E.K., Khouri H.M., Fedorova N.B.,
RA Carthy H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Haia R.-C., McClarty G., Rank R.G.,
RA Bayoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydomophila caviae (Chlamydia peitraci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147 (2003).
DR EMBL; AE016995; AAP05035.1; -; Genomic_DNA.
DR TIGR; CCA00284; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005546; Auto_transpbeta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR006626; Pbh1.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam PMP; 6.
DR SMART; SMO0710; Pbh1; 4.
DR TIGRFAMs; TIGR01414; autotrans barl; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 5.
KW Complete proteome.
SQ SEQUENCE 926 AA; 98213 MW; 0E1062E0778658B2 CRC64;

Alignment Scores:
Pred. No.: 3,6e-64 Length: 926
Score: 1099.00 Matches: 315
Percent Similarity: 43.8% Conservative: 154
Best Local Similarity: 29.4% Mismatches: 398
Query Match: 13.9% Indels: 204
DB: 2 Gaps: 33

US-10-701-844-1 (1-4435) x Q823W9_CHLCV (1-926)
QY 382 ATGCAACGCTCTCCCAATAGTTCTTTCTTCAATGATCTAGCTATTCTGTCGCTCT 441
DB 1 MetArgProSerLeuTyrLeuLeuLeuSerSerThrThrIlePro---LeuSer 19
QY 442 TTAATGGGGGGGATATGACAGCAAAATCATGTTCTCTCAA----- 483
DB 20 PheHisPheSerGlnValHisAlaGluValAlaLeuThrGlnGluSerValLeuAspAla 39
QY 484 ---GGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTACTATGTTATAGGAGAT 540
DB 40 AsnGlyAlaPheSerProGlnSerThrSerThrAla----- 51
QY 541 CGAGTGGGACTACTGTTTCTCGCAGAGAGTTAACTAAATAAATCTTGACAAATCT 600
DB 52 ---GlyGlyThrThrTyrAsnValGluSerAspIleSerIleValAspAlaGlyGlnAla 70
QY 601 ATTGACGCTTTCCTTAAAGTTGTTTGGGAACCTATTAGGAGTTTACTGTTTAGGG 660
DB 71 ---AlaAlaMetValSerAlaAlaPheValGlnThrAlaAspLeuThrPheLeuGly 89
QY 661 AGAGACACTGTTGACTTTCGAGAACATACGACTTCTACAAATGGGGAGCTCTAAGT 720
DB 90 AsnGlyAgsSerLeuAlaIleGluAsnValAsnSerGlyAlaAsnProGlyAlaIleTyr 109
QY 721 AATAGCGTCTGATGAGCTGTTTACTATTAGGGTTTAAAGAAATATCTTTTCCAAT 780
DB 110 ValSerAlaAlaAspLysThrLeuThrLeuThrAspPheSerThrLeuSerPheLysLys 129
QY 781 TGCAATTCATTCTTGCCGTACTGCTGCAACGACTAATAAGGGTAGCCAGACTCG 840
DB 130 Cys-----ProLysHisThrValAsnThrGly----- 138
QY 841 ACGAACACATCTACACCGTCTAATGCTACTATTATTTCTAAACAGATCTTTTGTATCTC 900
DB 139 -----LysGlyAlaValLysSerGlyGlyAlaLeuAsnLeuAla 151

```

```

QY 901 AATAATGAGAAGTCTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGAT 960
DB 152 AsnAsnAlaSerIleLeuPheAsnGlnAsnHisSerAlaGluAspGlyGlyAlaIleSer 171
QY 961 GCTAAGGAGTTAAACGGTTCAAGGAATTAACGAAGCTTTGTGTCTTCCCAAGAAATATGCT 1020
DB 172 CysLysAlaPheSerLeuThrGlySerSerLys----- 182
QY 1021 CAAGCTGATGGGAGCTTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAACGAGGCT 1080
DB 182 ----- 182
QY 1081 CCTATTGCCTTTGTAGCGAATTTGCAGAGTAGAAGAGGGGAGGATTCGCTGTTTCAG 1140
DB 183 GluIleSerPheThrThrAsnThrSerThrLysGlyGlyAlaIleAlaIleAlaThr--- 201
QY 1141 GATGGCGCAGCAGGAGTGTCTATCTTCAACAGAAGATCCAGTAGTAGTAAAGTTTTC 1200
DB 202 -----GlyValAlaAsnLeuSerAspAsnGlnGlyLysValIle---PheSer 216
QY 1201 AGAAATACGCGTAGAGTTTGTATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTTCC 1260
DB 217 GlyAsnThrAlaVal-----AsnSerGlyGlyAlaValTyrAla 229
QY 1261 TACGGGAACGTTGCTTCTCGAATAATAGGAACCTTGTTCCTCAACAATGTTGCTTCT 1320
DB 230 GluAlaAsnThrThrIleAlaGlyAsnSerAlaValValPheAsnAsnAlaVal--- 248
QY 1321 CCTGTTTACATTTGCTGCTAAGCAACCAAGTGGACAGGCTCTTAATACAGTAATAAT 1380
DB 249 -----ThrGlyThrThr 252
QY 1381 TACGAGATGGAGGAGCTATCTTCTGT---AAGAAATGTCGCGCAAGCAGGATCCAATAAC 1437
DB 253 AspGlyCysGlyGlyAlaIleHisCysSerLysThrGlyAla----- 266
QY 1438 TCTGATCAGTTTCTTTGATGGAGGGAGTAGTTTCTTTAGTAGCAATGATGCTGCT 1497
DB 267 ThrProValLeuThrIleArgAspAsnLysValLeuLeuPheLysGluAsnThrSerAla 286
QY 1498 GGGAAAGGGGAGCTATTTATGCCAAAAGCTTCGGTGTGTAACCTGTCGCCCTGTACAA 1557
DB 287 AlalysGlyGlyAlaIleTyrAlaAspLysLeuTyrLeuThrSerGlyGlyProThrVal 306
QY 1558 TTTTAAAGGAATATCGCTAATAGAT-----GGTGGAGCGATTTATTTAGGAGAA 1605
DB 307 PheValGlyAsnLysAlaThrAsnAlaAlaProLysGlyGlyAlaIleGlyIleAlaAla 326
QY 1606 TCTGAGAGCTCAGTTTATCTGCTGATATGAGATATTATTTTCGATGGGAATCTTAA 1665
DB 327 AsnGlyGluCysSerIleThrAlaGluHisGlyAspIleThrPheGluAsnAsnLeuIle 346
QY 1666 AGAACACCAAGAGAATGCTGCCGATTAATGGCGTAACGTGTCTCTCACAGCCAT 1725
DB 347 AlaThrAla-----AsnAsnAlaThrValLysArgAsnAlaIle 359
QY 1726 TCGATGGGATCGGGAGGAAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTC 1785
DB 360 AsnIleGluGlyAsnGlyLysPheValAsnLeuArgAlaAlaSerGlyAsnThrLeuThr 379
QY 1786 TTTAATGATCCATCGAGATGGCAACGGAATAACAGCCAGCGCAGCTTCCAAACTT 1845
DB 380 PheTyrAspProIleValValGly-----GlyThrAlaAlaAspLeu 393
QY 1846 CTAAAAATTAACGATGGTGAAGGA-----TACACAGGGGATATTGTTTTGCT--- 1893
DB 394 LeuThrLeuAsnGlnAlaGluGlyThrLysValTyrAsnGlyArgIleIlePheSerGly 413
QY 1894 -----AATGGAAGCAGTACTTTTGTACAA 1917
DB 414 GluLysLeuThrGluAspGlnThrAlaAspAlaAspAsnLeuLysThrValPheThrGln 433
QY 1918 AATGTTACGATAGCAGGAGGAGGATTTGTTCTTCTGTAAGGCAAAATATTCAGTGAAT 1977

```

```

Db      434  ProfileAlaLeuAlaAlaGlyGluLeuIleLeuAsgAsnGlyValGluValAlaLys 453
QY      1978  TCTTAAGTCACAGAGTGGAGTCTGTATATG---GAAGCTGGAGATCATCGGATTTT 2034
Db      454  AlaValSerGlnThrAlaGlySerLeuIleLeuIleAspAlaGlyThr----- 469
QY      2035  GTAACCTCCACACACCACAGAGCTCTCGCGCTTAATCACTGATCATCGCTTCCCAAT 2094
Db      470  -----LysLeuSerAlaLysThrGluAspValThrLeuThrAsn 482
QY      2095  CTGCATTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2154
Db      483  LeuAlaIleAsnProAsnSerLeu----- 490
QY      2155  CCTCAGCGCAAGATCTCATCTCGAGTCATTTGGTAGC---ACAATGCTCGTCTGTT 2211
Db      491  -----AspGlyThrLysIleAlaValIleAlaAlaValAlaAsnAlaLysAsnVal 507
QY      2212  ACAATTAGTGGCGCTATCTTTTGGAGATTTGGATGATACAGCTTATGATAGTATGAT 2271
Db      508  ThrValThrGlyAlaIleGlyIleIleAspProThrGlyLysPheThrGluAspHisLys 527
QY      2272  TGGTAGGTCTTAATCAAAATCAATGTCCTGMAATTACAGTTAGGACTAAG----- 2325
Db      528  -----LeuAsnGluThrLeuAlaLeuGlyIleGlnPheSerAlaLysGlySer 544
QY      2326  ---CCCCAGCTAATGCCCATCGATTTGACTCTTAGGGAATGAGATGCTAAGTATGCG 2382
Db      545  IleThrThrThrAspValProSer---ThrThrThrArgSerProAlaGlnHisThrGly 563
QY      2383  TATCAAGGAAGCTGGAGCTTGGCTGG-----GATCCTAATACAGCA 2424
Db      564  TyrGlnGlyAsnTrpSerLeuSerTrpIleThrAspAsnGlySerAspProLysThr--- 582
QY      2425  AATAATGGTCTTATCTCTGAAAGCTACATGACTGAAACTGGGTATATATCTCGGCT 2484
Db      583  -----GlnThrAlaValPheAsnTrpAsnLysThrGlyThrAsnProAsnPro 598
QY      2485  GAGCGAGTAGTCTTTTGGTCCAAATAGTTTATGGGATGCCATTTTAGATATACATCT 2544
Db      599  GluArgAlaProLeuValLeuAsnSerLeuTrpGlySerPheMetAspIleArgSer 618
QY      2545  GGCATTACAGCAATTCAGCAAGTGTGGATGGG---CGCTCTTATTGTCGAGGATTTAG 2601
Db      619  IleGlnAspValMetGluArgSerValAspThrLeuLeuGluThrArgArgGlyLeuTrp 638
QY      2602  GTTCTCGAGTTTCGAATTTCTCTATCATGACCGCGATGCTTTTAGTCAGGGATATCGG 2661
Db      639  ValSerGlyValGlyAsnPheLeuHisLysAspProSerAlaGluAsnArgLysPheArg 658
QY      2662  TATATTAGTGGGGTATTCTCTTAGGAGCAAACTCCTACTTTTGG---TCATCGATGTTT 2718
Db      659  HisIleSerSerGlyThrValLeuLeuGlyAlaThrThrAsnThrSerGlnGluAspThrLeu 678
QY      2719  GGTCTAGCATTTACCGAAGTATTGTTAGATCATCTAAGATTATGATGTCGCTCTTCCAAT 2778
Db      679  SerValAlaPheCysGlnLeuPheGlyLysAspLysAspThrLeuValAlaLysAsnAla 698
QY      2779  CATCATGCTGTCATAGGATCCGTTTAT----- 2805
Db      699  AlaAsnValThrAlaGlySerIleThrGlnHisValSerLysPheAspAspLeuThr 718
QY      2806  -----CTACTCTACCCCAACAGCTTTTATGTGGATCTCTAT-----TTGTTC 2844
Db      719  ArgLeuPheAsnGlyProAsnThrCysCysSerGlyPheSerLysGluIleProIlePhe 738
QY      2845  GAGATCGGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATAT 2904
Db      739  LeuAspAla-----GlnValThrThrCysHisThrAsnAsnAsnMetThrThrThrThr 756
QY      2905  ACATTCAGAGGAGGAGCGATGTTCTGTTGGATATTAATCTGCTGGCTGGAGAGATTGGA 2964

```

```

Db      757  ThrAspTyrProGluValLysGlySerTrpGlyAsnAspThrValGlyValAlaLeuSer 776
QY      2965  GCGGATTTACCGATTTGTGATTACTTCAATCAAGCTCTATTTGAATGAGTTCGCTCTTTC 3024
Db      777  ThrSerValProIleProIlePhe---ThrHisAlaPhePheAspSerTyrAlaProPhe 795
QY      3025  GTCAAGCTGAGTTTCTTATGTCGCGATCATCAATCTTTTACAGAGGAAGCGGATCACT 3084
Db      796  AlaLysLeuGlnValValTyrAlaHisGlnGluAspPheLysGluProThrArgGluGly 815
QY      3085  CGGCGATTCAGAGCGGACATCTCTTAATCTATCACTGCTCTGTTGGAGTGAAGTTTGTAT 3144
Db      816  ArgThrPheGluSerSerAspLeuAsnValSerValProIleGlyValLysPheGlu 835
QY      3145  CGATGTTCTAGTACACATCTTAATAATATAGCTTTATGGCGCTTATCTGTGATGCT 3204
Db      836  LysLeuValTyrGlyGluLysThrAlaTyrAspLeuThrLeuMetTyrValProAspVal 855
QY      3205  TATCGCACCATCTCTGTGTAAGTACAGACACGCTCTCTATCCCATCAAGACATGGACACA 3264
Db      856  TyrArgHisAsnProAsnCysIleThrGlyPheAlaIleAsnAspValThrTrpLeuThr 875
QY      3265  GATGCTCTTATTTAGCAAGACATGGAGTTTGGTTAGAGGATCTATGTATGCTTCTCTA 3324
Db      876  ThrAlaThrAsnLeuAlaArgGlnAlaPheIleIleArgAlaGlyAsnHisIleAlaVal 895
QY      3325  ACAAGTATATAGAGTATATGGCCATATGAGATATGATATCGAGATGCTTCTCGAGGC 3384
Db      896  ThrSerGlyPheGluMetPheSerGlnPheGlyPheGluLeuArgSerSerArgAsn 915
QY      3385  TATGTTTGTAGTGCAGGAAGTAGATCCGGTTC 3417
Db      916  TyrAsnValAspLeuGlyAlaLysValSerPhe 926

RESULT 23
PMP11 CHLPN
ID PMP11 CHLPN STANDARD; PRT; 928 AA.
AC O86164; O9K29;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmp11 precursor (Polymorphic membrane
DE protein 11) (Outer membrane protein 4).
GN Name=pmp11; Synonyms=omp4; OrderedLocusNames=CpN0449, Cp0302, Cp0468;
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029/VR-1310;
RX MEDLINE=99081766; PubMed=9864239;
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity."
RL Am. Heart J. 138:S491-S495(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
RX MEDLINE=99026606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
RN [4]

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=AR39; PubMed=10684935; DOI=10.1093/nar/28.6.1397;  
 RA MEDLINE=20150255; Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J.D., Uterback T.R., Berry K.J.,  
 RA Bass S., Linher K.D., Weidman J.P., Khouri H.M., Craven B., Bowman C.,  
 RA Dodson R.J., Gwin M.L., Nelson W.C., DeBoy R.T., Kolonay J.P.,  
 RA McClary G., Salzberg S.L., Eisen J.A., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
 RT pneumoniae AR39.";  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=J138;  
 RA MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;  
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishi K., Hattori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 RT from Japan and CML029 from USA.";  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=TW-183;  
 RA Geng M.M., Schumacher A., Muehldorfer I., Benach K.W., Schaefer K.P.,  
 RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;  
 RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with  
 RT other Chlamydia strains based on whole genome sequence analysis.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)  
 CC (Potential).  
 CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; AJ001311; CA04672.1; -; Genomic DNA.  
 CC EMBL; AJ133034; CAB37072.1; -; Genomic DNA.  
 CC EMBL; AB001628; AAD18593.1; -; Genomic DNA.  
 CC EMBL; AE002192; AAF38159.1; ALT INIT; Genomic DNA.  
 CC EMBL; BA000008; BAA98658.1; -; Genomic DNA.  
 CC EMBL; AE017158; AAP98399.1; -; Genomic DNA.  
 CC PIR; D72077; D72077.  
 CC PIR; H86546; H86546.  
 CC PHCI-2DPAGE; O86164; -.  
 CC TIGR; CP0302; -.  
 CC InterPro; IPR005546; Auto transpbeta.  
 CC InterPro; IPR011427; ChlamPMP\_M.  
 CC InterPro; IPR003368; Chlamydia PMP.  
 CC Pfam; PF03797; Autotransporter; 1.  
 CC Pfam; PF02415; Chlam PMP; 6.  
 CC Pfam; PF07548; ChlamPMP\_M; 1.  
 CC TIGRfam; TIGR01376; POMP\_repeat; 5.  
 KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.  
 FT SIGNAL  
 FT CHAIN 1 24 Potential.  
 FT CHAIN 25 928 Probable outer membrane protein pmp11.  
 SQ SEQUENCE 928 AA; 98904 MW; 789BCDD62C911402 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.43e-63 Length: 928  
 Score: 1090.00 Matches: 311  
 Percent Similarity: 43.0% Conservative: 154  
 Best Local Similarity: 28.8% Mismatches: 394  
 Query Match: 13.8% Indels: 222  
 DB: 1 Gaps: 31  
 US-10-701-844-1 (1-4435) x PMP11\_CHLPPN (1-928)  
 QY 382 ATGCAAGCTCTTCCATAGTCTCTTCTTCAATGATCTAGCTATTCTTGC----- 435  
 DB 1 MetLysThrSerIleProTrpValLeuValSerSerValLeuAlaPheSerCysHisLeu 20

QY 436 TGCTCTTTAAATGGGGGGGATATGACGAGAAATCATGGTTCCTCAAGAAATTTACGAT 495  
 DB |||||  
 DB 21 GlnSerLeuAlaAan-----GluGluLeuLeuSerProAspAspSerPheAan 36  
 QY 496 GGGGAGAGCTTAACCTGTATCATTT-----CCCTATCATGTTATA 534  
 DB |||||  
 DB 37 GlyAsnIleAspSerGlyThrPheThrProLysThrSerAlaThrThrTyrSerLeuThr 56  
 QY 535 GGAGATCCGAGTGGGACTACTGTTTTTCTGCAGGAGAGTTAAACATTAATAAAATCTTGAC 594  
 DB |||||  
 DB 57 GlyAsp-----ValPhePheTyrGluProGlyLysGlyThr----- 68  
 QY 595 AATTCATATGAGCTTTGCCCTTTA-----AGTTGTTTGGGAACTTATATAGGAGTTTT 648  
 DB |||||  
 DB 69 -----ProLeuSerAspSerCysPheLysGlnThrThrAspAsnLeu 82  
 QY 649 ACTGTTTTAGGAGAGACACTCGTTCGACTTTCGAGAACATACGAGCTTCTACAAATGGG 708  
 DB |||||  
 DB 83 ThrPheLeuGlyAsnGlyHisSerLeuThrPheGlyPheIleAspAlaGlyThrHisAla 102  
 QY 709 GCAGCTCTAAGTAAATAGCGTCTGCTGAGCTGTTTACTATTGAGGGTTTTTAAAGAATTA 768  
 DB |||||  
 DB 103 GlyAlaAlaAlaSerThrThrAlaAsnLysAsnLeuThrPheSerGlyPheSerLeuLeu 122  
 QY 769 TCCTTTTCCAAATTCGAATTCATTTACTTCCGCTACTGCTGCTGCAACGACTTAATAAGGT 828  
 DB |||||  
 DB 123 SerPheAspSer----- 126  
 QY 829 AGCCAGACTCCGACGACAAACATCTACACCGTCTAATGCTACTATTATTCTTAAACAGAT 888  
 DB |||||  
 DB 127 -----SerProSerThrThrValThrThrGlyGlnGlyThrLeuSerSerAlaGlyGly 144  
 QY 889 TTTTGTGTACTCAATAATGAGAGTCTCTCATTTCTATAGTAATTTAGTCTCGGAGATGGG 948  
 DB |||||  
 DB 145 ValAsnLeuGluAsnIleAArgLysLeuValValAlaGlyAsnPheSerThrAlaAspGly 164  
 QY 949 GGAGCTATAGATGCTAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAA 1008  
 DB |||||  
 DB 165 GlyAlaIleLysGlyAlaSerPheLeuLeuThrGlyThrSerGlyAspAlaLeuPheSer 184  
 QY 1009 GAAATACTGCTCAAGCTGATGGGGAGCTGTCAAGTAGTACGACAGTTTCTGCTATG 1068  
 DB |||||  
 DB 185 AsnAsnSerSerSerThrLysGlyGlyAla----- 194  
 QY 1069 GCTAACGAGCTCTCTATTGCTTTGCTAGCGAATGTTGACGAGTAAGAGGGGAGGAT 1128  
 DB |||||  
 DB 195 -----IleAlaThrThrAlaGlyAlaArg----- 202  
 QY 1129 GCTGCTGTTACGAGTGGGACGAGGAGTGTCTATCTTCTCAACAGAGATCCAGTA 1188  
 DB ----- 202  
 QY 1189 GTAAGTTTTTCCAGAAATACTCGG---GTAGAGTTTGTATGGGAACGTAGCCGAGTAGGA 1245  
 DB |||||  
 DB 203 -----IleAlaAsnAsnThrGlyTyrValArgPheLeuSerAsnIleAlaSerThrSer 220  
 QY 1246 GGAGGAGATTTACTCTCTACGGGAACGTTGCTTCTGAATATATGAAAAACCTTTGTTCTC 1305  
 DB |||||  
 DB 221 GlyGlyAlaIleAspAspGluGlyThrSerIleLeuSerAsnAsnLysPheLeuTyrPhe 240  
 QY 1306 AACAAATGTTGCTTCTCTGTTTACATTGCTGCTAAGCAACCAACAGTGGGACAGCTTCT 1365  
 DB -----GluGlyAsnAlaAla 245  
 QY 1366 AATACGAGTAATAATTACGGAGATGGAGAGCTATCTTCTGTAAGAATGGTGGCAAGCA 1425  
 DB |||||  
 DB 246 LysThrThr-----GlyGlyAlaIle-----CysAsnThrLysAlaSerGly 259  
 QY 1426 GGA-----TCCAAATCACTCTGGATCATGTTCTTCTTGATGGAGAGGAGTA 1470  
 DB |||||  
 DB 260 SerProGluLeuIleIleSerAsnAsnLys-----Thr 270





ID AC Q92393, Q9RB66; STANDARD; PRT; 930 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Probable outer membrane protein pmp8 precursor (Polymorphic membrane  
 DE protein 8) (Outer membrane protein 11)  
 GN Name=pmp8; Synonyms=omp11; OrderedLocNames=CpB0446, CP0307, CpB0463;  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae)  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CWL029 / VR1310;  
 RX MEDLINE=20007584; PubMed=10539856;  
 RA Christiansen G., Bosen T., Hjerno K., Daugeard L., Mygind P.,  
 RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;  
 RT "Molecular biology of Chlamydia pneumoniae surface proteins and their  
 RT role in immunopathogenicity";  
 RL Am. Heart J. 138:S491-S495(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=CWL029;  
 RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;  
 RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";  
 RL Nat. Genet. 21:385-389(1999).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J.D., Uitterback T.R., Berry K.J.,  
 RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,  
 RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,  
 RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.; Chlamydia  
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
 RT pneumoniae AR39";  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;  
 RA Shirai M., Hirakawa H., Kinoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 RT from Japan and CWL029 from USA";  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=TW-183;  
 RA Geng M.M., Schumacher A., Muehlhofer I., Bensch K.W., Schaefer K.P.,  
 RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;  
 RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with  
 RT other Chlamydia strains based on whole genome sequence analysis";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)  
 CC (Potential).  
 CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.  
 CC -!- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL; AJ133034; CAB37068.1; -; Genomic DNA.  
 CC EMBL; AB001627; AAD18590.1; -; Genomic DNA.  
 CC EMBL; AB002193; AAF38164.1; -; Genomic DNA.  
 CC EMBL; BA000008; BAA98654.1; -; Genomic DNA.  
 CC EMBL; AE017158; AAP98394.1; -; Genomic DNA.

DR PIR; A81591; A81591.  
 DR PIR; D72078; D72078.  
 DR PHCI-2DPAGS; Q92393; -.  
 DR TIGR; CP0307; -.  
 DR InterPro; IPR005546; Auto transportbeta.  
 DR InterPro; IPR006315; Auto transporter.  
 DR InterPro; IPR011427; ChlamPMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 6.  
 DR Pfam; PF07548; ChlamPMP\_M; 1.  
 DR TIGRFAMs; TIGR01414; autotrans barl; 1.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 6.  
 KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.  
 FT SIGNAL 26 Potential.  
 FT CHAIN 27 930 Probable outer membrane protein pmp8.  
 FT VARIANT 177 177 T -> A (in strain AR39, strain J138 and  
 FT strain TW-183).  
 SQ SEQUENCE 930 AA; 97670 MW; 46A9B5E3BB913C4C CRC64;  
 Alignment Scores:  
 Pred. NO.: 7.68e-63 Length: 930  
 Score: 1079.00 Matches: 303  
 Percent Similarity: 43.2% Conservative: 160  
 Best Local Similarity: 28.3% Mismatches: 407  
 Query Match: 13.7% Indels: 202  
 DB: 1 Gaps: 29  
 US-10-701-844-1 (1-4435) x PMP8\_CHLPN (1-930)  
 QY 382 ATGCAAGCGTCTTCGATTAAGTCTTCTTCAATGATTTAGCTATTCTTGTGCTCT 441  
 Db 1 MetLysileProLeuHisLeuLeuSerThrLeuVal---ThrProileu 19  
 QY 442 TTAATGGGGGGGATATGACGACGAAATC---ATGTTCTCTCAAGGAATTTACGATGGG 498  
 Db 20 LeuSerilealaThrTyGlyAlaAspAlaSerLeuSerProThrAspSerPheAspGly 39  
 QY 499 GAGACGTTAATCTGATTCATTTCCCTATCTGTTTATAGGAGTCGACGATGCTGTT 558  
 Db 40 AlaGlyGlySerThrPheThrProLysSerThr---AlaAspAlaAsnGlyThrAsnTyr 58  
 QY 559 TTTCTGACGAGGAGTTAACTATAAATAATCTTCAATTTCTATTGACGCTTGCCTTTA 618  
 Db 59 ValLeuSerGlyAsnValTyIleAsnAspAlaGlyLysGly---ThrAlaLeuThrGly 77  
 QY 619 AGTTGTTTGGAACTTATTAGGAGTTTACTGTTTATAGGAGGAGGACACTCGTCTGACT 678  
 Db 78 CysCysPheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTyrSerPheSer 97  
 QY 679 TTCGAGAACATACGACTTCTACAAATGGGGGAGCTCTAAGTAATAGCGCTGCTGATGGA 738  
 Db 98 PheAsnThrValAspAlaGlySerAsnAlaGlyAlaAlaAlaSerThrThrAlaAspLys 117  
 QY 739 CTGTTTACTATTAGGGGTTTAAAGATTTATCTTTTCCAAATGCAATTCATTTACTTCC 798  
 Db 118 AlaLeuThrPheThrGlyPheSerAsnLeuSerPhe-----Ile 130  
 QY 799 GTACTGCTGCTGCAACGACTATAAGGTTAGCGAGCTCCGACGACACACA----- 849  
 Db 131 AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu 150  
 QY 850 TCTACACGGCTAATAGTACTATTATTCTTAAACACAGATCTTTTGTCTACTCAATTAATGAG 909  
 Db 151 AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal----- 164  
 QY 910 AAGTTCTCTATCTATCTAATTTAGTCTCTGGAGATGGGGGAGCTATAGATGCTAAGACG 969  
 Db 165 -----SerAsnGluAlaAsnAsnAsnGlyGlyAlaIleThrThrLysThr 179  
 QY 970 TTAACGGTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAGAGAAATATCTGCTCAAGCTGAT 1029  
 Db 180 LeuSerileSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLeuLeu 199

QY 1030 GGGGAGCTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTTAACGAGCTCCTATTGCC 1089  
DB 200 GlyGlyAlaIleTyrSerSerAlaAlaAlaSerIleSerGlyAsnThrGlyGlnLeuVal 219  
QY 1090 TTTGTAGCGAATGTTGCAGGAGTAGAGAGGGGAGGAGTTCGTCTGTTCAGGATGGCGAG 1149  
DB 220 PheMetAsnAsn---LysGlyGluThrGlyGlyGlyAlaLeuGlyPheGlu----- 235  
QY 1150 CAGGAGGTGTCATCATCTACTCTCAACAGAGATCCAGTAGTAGTATTTTCCAGAAATACT 1209  
DB 236 -----AlaSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr 252  
QY 1210 GCGGTAGAGTTTGTAGGGAGCGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC 1269  
DB 253 AlaThrAsp----- 255  
QY 1270 GTTGCTTTCTGTAATAATGGAATAACCTGTGTTCTCAACAATGTTGCTTCTCCTGTTTAC 1329  
DB 255 ----- 255  
QY 1330 ATTGCTGTCTAAGCAACCAAGTGGACAGGCTTCTAATACGAGTAATAATTACGAGAT 1389  
DB 256 ---AlaAla-----GlyLys 259  
QY 1390 GAGGAGCTATCTCTGTGAAGAAATGTCGCGAAGCGAGATCCAAATCACTCTGATCAGTT 1449  
DB 260 GlyGlyAlaIleTyrCysGlyLys-----ThrGlyGluThrProThrLeu 274  
QY 1450 TCTTTGATGGAGGAGGTAGTTTCTTTAGTACCAATGTAGCTGCTGGGAAGGGGA 1509  
DB 275 ThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThrGlnGly 294  
QY 1510 GCTATTATTGCCAAAAGCTCTCGGTTGCTTAACCTGTGGCCCTGTACAAATTTTAAAGAAAT 1569  
DB 295 AlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn 314  
QY 1570 -----ATCGCTAATGATGTGTGGAGCGCAATTTATTAGGAGAACTCGAGAG 1614  
DB 315 ArgCysGlyAsnThrAlaAlaGlyLysGlyGlyAlaIleAlaIleAlaAspSerGlySer 334  
QY 1615 CTCAGTTTATCTGCTATGAGATATATTTTCGATGGGAATCTTAAAGAACAGCC 1674  
DB 335 LeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsn----- 349  
QY 1675 AAAGAGAATGCTCCGATGTTAATGGCTAACTGTCTCTCACAGCCATTTTCGATGGGA 1734  
DB 350 -----ThrLeuThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyrLeuGly 367  
QY 1735 TCGGGAGGGAATAACGACATTAAAGAGCTAAAGCAGGCGCATCAGATCTCTTTAATGAT 1794  
DB 368 SerSerAlaLysIleThrAsnLeuAlaAlaGlnGlyGlnSerIleTyrPheTyrAsp 387  
QY 1795 CCATCGAGATGCCAACGGAAT-----AACCAGCCAGCGCAG 1833  
DB 388 ProIleAlaSerAsnThrThrGlyAlaSerAspValLeuThrIleAsnGlnProAspSer 407  
QY 1834 TCTTCCAACTCTTAAAAATTACGATGGTGAAGGATACACAGGGGATATTTGTTTGTCT 1893  
DB 408 AsnSerProLeu-----AspTyrSerGlyThrIleValPheSer 420  
QY 1894 -----AATGAACGATCTTTGTATC 1914  
DB 421 GlyGlyLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys 440  
QY 1915 CAATAATGTTAGATAGACGAGGAGGAGTGTCTTCTCGTGAAGGCAAAATATCATGTC 1974  
DB 441 GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLeuLysGlyAsnValGluLeuAspVal 460  
QY 1975 AATTCTCTAAGTCAGACAGGTGGAGTCTGTATATGAAGCTGGGAGTACATGGGATTTT 2034  
DB 461 AsnGlyPheThrGlnThrGluGlySerThrLeuLeu----- 472

QY 2035 GTAACCTCCACCAACCAACACAGCTCTCTGCCCTAATACAGTTGTATCACGCTTTCCAAAT 2094  
DB 473 -----MetGlnProGlyThrLysLeuLysAlaAspThrGlyAlaIleSerLeuThrLys 490  
QY 2095 CTGCAATTTGCTCTCTTTCTTTGTTAGCAACAATGCAGTTACGAATCCTCTCCTCAACAT 2154  
DB 491 LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer----- 505  
QY 2155 CCTCAGCGCAAGATTCTCATCTCTGCAGTCATTGCTAGCACAACTGCTGGT----- 2205  
DB 506 -----IleGluThrAlaGlyAlaAsnLys 513  
QY 2206 TCTGTTCAATTAGTGGCCTATCTTTTTCAGATTTGGAGTACAGCTTATCATAGG 2265  
DB 514 ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyrGlu--- 532  
QY 2266 TATGATGGCTAGGTTCTAATCAAAAATCAATGTCCTCTGAAATTCACATAGGAGCTAAG 2325  
DB 533 -----SerHisThrIleAsn-----GlnAlaPheThrGln 542  
QY 2326 CCC-----CCAGCTAATGCCCATCAGATTGACTCTAGGG----- 2361  
DB 543 ProLeuValValPheThrAlaAlaThrAlaAlaSerAspIleTyrIleAspAlaLeuLeu 562  
QY 2362 -----AATGAGATGCTAAGTATGCTATCAAGAAAGCTGGAAGCTTGGC 2406  
DB 563 ThrSerProValGlnThrProGluProHisTyrGlyTyrGlnGlyHisTrpGluAlaThr 582  
QY 2407 TGG-----GATCTTAATACAGCAAAATATGTCCTCTTATCTCTGAAAGCTACATGACTAAA 2463  
DB 583 TrpAlaAspThrSerThrAlaLysSerGly-----ThrMetThrTrpValThr 598  
QY 2464 ACTGGGTATATCTCGGCTGAGGAGTAGCTCTTTGGTTCCAAATAGTTTATGGGGA 2523  
DB 599 ThrGlyTyrAsnProAsnProGluArgAlaSerValValProAspSerLeuTrpAla 618  
QY 2524 TCCATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGGCGCTCT 2583  
DB 619 SerPheThrAspIleArgThrLeuGlnIleMetThrSerGlnAlaAsnSerIleTyr 638  
QY 2584 TATGTCGAGGATATGCGTTTTCGAGTTTCGAATTTCTTCTATCATCATCCGCGATGCT 2643  
DB 639 GlnGlnArgGlyLeuTrpAlaSerGlyThrAlaAsnPhePheHisLysAspLysSerGly 658  
QY 2644 TTAGTCTCAGGATATCGGTATATAGTGGGGTTATTCCTTAGGCAACACTCTAC--- 2700  
DB 659 ThrAsnGlnAlaPheArgHisLysSerTyrGlyTyrIleValGlySerAlaGluAsp 678  
QY 2701 TTTGGATCATCATGTTTGGCTTAGCATTTACCGAGTATTTGGTAGATCTAAAGATTAT 2760  
DB 679 PheSerGluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysAspLeu 698  
QY 2761 GTAGTGTGTCGTTCCAATCATCATGCTTCATAGGATCGGTTTATCTATCTACCCCAAA 2820  
DB 699 PheIleValGluAsnThrSerHisAsnTyrLeuAlaSerLeuTyrLeuGlnHisArgAla 718  
QY 2821 GCTTTATGTCGATCCTATTG-----TTCCGAGATGCGGTT----- 2856  
DB 719 PheLeuGlyGlyLeuProMetProSerPheGlySerIleThrAspMetLeuLysAspIle 738  
QY 2857 -----ATCGCTGCTAGCTACGGGTTTGGGAATCAGCATATGAAAACCTCA 2901  
DB 739 ProLeuIleLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThrArg 758  
QY 2902 TATACATTTTCAGAGGAGGAGCGATGTTGCTGGGATTAATAACTGTCTGGCTGGAGAGATT 2961  
DB 759 TyrThrSerTyrProGluAlaGlnGlySerTrpThrAsnAsnSerGlyAlaLeuGluLeu 778  
QY 2962 GAGCGGGATTACCGATTGATGATTCTCCATCTAAGCTCTATTGATGAGTTGCGCTCT 3021  
DB 779 GlyGlySerLeuAlaLeuTyrLeuProLysGluAlaProPhePheGlnGlyTyrPhePro 798  
QY 3022 TTCGTGCAAGCTGAGTTTCTTATGCGATCATGAATCTTTTACAGAGGAGCGCATCAA 3081

```

Db 799 PheLeuIysPheGlnAlaValTyrSerArgGlnGlnAsnPheLysGluSerGlyAlaGlu 818
Qy 3082 GCTCGGGCAATCAAGAGCGGACATCTCTAAATCTATCATGTTCTCTGTTGGAGTGAAGTTT 3141
Db 819 AlaArgAlaPheAspGlyAspLeuValAAsnCysSerIleProValGlyIleArgLeu 838
Qy 3142 GATCGATGTTCTAGTACACATCTCTAAATATATAGCTTTATGCGGCTTATATCTGTGAT 3201
Db 839 GluIySerGluAspGluLysAsnAAsnPheGluIleSerLeuAlaTyrIleGlyAsp 858
Qy 3202 GCTTATCCACCATCTCTGCTAGTACAGCAAGCTCTCTATCCATCAAGACATGGACA 3261
Db 859 ValTyrArgLysAsnProArgSerArgThrSerLeuMetValSerGlyAlaSerTyrPhe 878
Qy 3262 ACAGATGCTTTTCATTAGCAAGACATGAGATGTTGTGTTAGAGGATCTATGATGCTTCT 3321
Db 879 SerLeuCysLysAsnLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeuThr 898
Qy 3322 CTAACAGTATATAGATATATGCGCATATGCGCATATGATGATGATGATGATGATGATGATG 3381
Db 899 LeuSerProHisValGluLeuSerGlyGluAlaTyrGluLeuArgGlySerAlaHis 918
Qy 3382 GCCTATGTTTGGTGCAGGAGTAGAGTCCGGTTC 3417
Db 919 IleTyrAsnValAspCysGlyLeuArgTyrSerPhe 930

RESULT 25
Q5L6J2_CHLAB PRELIMINARY; PRT; 1378 AA.
AC Q5L6J2
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polymorphic outer membrane protein.
GN Name=omp15G; Synonyms=pomp145A; OrderedLocuNames=CAB283;
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiophila.
OX NCBI_TaxID=83555;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S26/3;
EX PubMed=15837807; DOI=10.1101/gr.3684805;
RA Livingston M., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Quail M.A., Price C., Barrett B.G., Parkhill J., Longbottom D.,
RA "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640(2005).
DR EMBL; CR48038; CAH63733.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 1378 AA; 144965 MW; 40B2108E7F26BP89 CRC64;

Alignment Scores:
Fred. No.: 1.16e-62 Length: 1378
Score: 1077.00 Matches: 330
Percent Similarity: 41.1% Conservative: 178
Best Local Similarity: 26.7% Mismatches: 382
Query Match: 13.7% Indels: 346
DB: 2 Gaps: 42

US-10-701-844-1 (1-4435) x Q5L6J2_CHLAB (1-1378)
Qy 499 GAGAGCTTAAGTATGATCTATCTCCCTATATCTGTTATA----- 534
Db 226 GluThrLeuGluSerSerLeuProHisThrProLeuSerPheThrAlaSerSerAspVal 245
Qy 535 -----GGAGATCCAGTGGGACTACTGTTTCTTCGAGGAGAGTTAACATTA 582
Db 246 AlaAsnGluHisGlySerProSerProGluThr----- 256
Qy 593 AAAAATCTTGACAAATCTATTGCGAGCT-----TTGCCCTTTAAGTTGTTTT 627

```

```

Db 257 LysThrProThrAlaValAlaSerGlnAlaAsnGlyGluThrAsnLysSerCysPhe 276
Qy 628 GGGAACTATTAGGAGTGTCTTACTGTTTATAGGAGAGGACACATCGTTGACTTTTCGAGAAC 687
Db 277 SerAsnThrAspGlySerLeuThrPheValGlyGlyAsnHisSerLeuThrPheSerAsn 296
Qy 688 ATACGGACTTCTCAAAATGGGAGCTCTAAGTAATAGCGCTGCTGTAGTGGACTGTTTACT 747
Db 297 IleSerValThrAlaProGlySerAlaIleAsnAsnSerAlaGlySerAlaLeu---Thr 315
Qy 748 ATGAGGCTTTTAAAGAAATATCTCTT-----TCCAAATGC 783
Db 316 PheSerGlyPheLysAspLeuSerPheValCysAlaThrAsnGlnAspGlnThrLysAla 335
Qy 784 AATTCAATCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
Db 336 AspSerAlaIleTyrValGlyProLysAlaThrSerGlyAspGlyGluLeuSerValGln 355
Qy 844 ACAACATCTACACGCTCTAATGCTACTATTATTCTTAAACAGATCTTTGTTTACTCAAT 903
Db 356 AlaGluGlySerLysAlaAsnSerHisThrGluGlySerValAsnThrAlaLeuIleLys 375
Qy 904 AATGAGAAGTCTCATCTATAGTAAATTAGTCTCTCGAGATGGGGAGCTATAGATGCT 963
Db 376 -----TyrAlaGly---IleProAlaSerAlaThrAlaThrAsnGly 388
Qy 964 AAGAGCTTAAACGGTTCAAGGAATTTAGCAAGCTTTGTGTCTTCCAAAGAA----- 1011
Db 389 GluSer-----GlnGluAspProAlaThr 396
Qy 1012 AATCTCTCAAGCTGATGGGAGGCTGTCAAGTAGTACACAGTTCCTGCTGCTGCTGCTGCTG 1065
Db 397 AsnValAlaProAlaThrGlnGlyGlyAlaGluLysGluGluAlaAsnSerProAsp 416
Qy 1066 -----ATGGCTAACGAGGCTCTTATGCTCTTGTAGCGAATGTTGCGAGGAGTAAAGAGG 1119
Db 417 IleIleLysGlnAsnValAsnIleThrPheSerSerAsnSerSerLysThrAlaGly 436
Qy 1120 GGAGGAGTGTGCTGTGTGCTGAGGATGGGAGGAGTGTGCTATCTATCTTCAACAGAA 1179
Db 437 GlyAlaIleArgVal-----SerGlySerAlaThrIleGluAsnThrGlyThr--- 453
Qy 1180 GATCCAGTAGTAAAGTTTTCAGAAATCTGCGGTAGAGTTGATGGAAAGTATGCGCGA 1239
Db 454 -----CysThrPheSerAsnAsnAlaLysGlu----- 463
Qy 1240 GTAGGAGGAGGATTTACTCTCTAGGGAAGTGTCTTCTCTGAATAATGAAAAACCTTG 1299
Db 464 GlnGlyAlaIleSerValAsnGlyAsnCysAspIleThrGlyAsnLysAsnValVal 483
Qy 1300 TTTCTCAACAAATGTTGCTCTCTCTGTT-----TACATT 1332
Db 484 PheSerGlyAsnGlnAlaGlnGluIleProAlaProSerIleValThrValGluAla 503
Qy 1333 GCTCTAAGCAACCAACAGTGGACAGGCTTCTTAACGAGTAAATAATTACGAGATGGA 1392
Db 504 AlaValGluGluProValGluAlaLys-----GlySerGly 515
Qy 1393 GGAGCTATCTCTGT----- 1407
Db 516 GlyAlaIleTyrCysValLysAlaProIleSerIleProAlaThrProIleGlnPro 535
Qy 1408 -----AAGAAATGGTCCGCAAGCAGGATCCAAATACTCT----- 1440
Db 536 LysGlnThrLeuProSerLeuLysSerAlaAlaLeuSerGlyGluThrHisAlaLysVal 555
Qy 1441 -----GGATCAGTTCTCTTGTATGAGGAGGAGTGTAGTT 1473
Db 556 AlaGlnLysGluGluProSerProAspProCysLeuThrIleSerGlyAsnAlaSerVal 575
Qy 1474 TTTCTTATGATGCAATGATGCTGCTGGGAAAGGGAGCTATTATGCAAAAGCTCTCG 1533

```

576 IlePheAspAsnAsnSerThrValThrGlyGlyAlaIleHisAlaLysLysValVal 595  
Db  
QY 1534 GTTGCTAACTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGATT 1593  
Db  
QY 596 LeuSerSerGlyAsnMetThrPheSerAsnAsnSerSerGlyGlyGlyAlaIle 615  
Db  
QY 1594 TATTTAGGAGACTCGGAGAGCTCAGTTTATCTGCTGATTAAGGAGATATATTTTCGAT 1653  
Db  
QY 616 TyrIleAlaAspGlyGlyAspIleSerIleThrAlaThrGlySerIleIlePheGln 635  
QY 1654 GGGAACTTTAAAGAACAGCCAAAGAGAA 1683  
Db  
QY 636 GlyAsn---LysValThrAlaAlaAspSerIleThrLeuProThrLysLysGluIleAla 654  
QY 1684 -----GCTGCCGATGTTAATCGGCTAAGTCTGCTCACAAGCC----- 1722  
Db  
QY 655 IleAlaAlaGluSerIleGluGlyThrAspProSerGlnAlaSerGlySerAlaSer 674  
QY 1722 ----- 1722  
Db  
QY 675 AlaGlyMetThrSerAlaPheThrLeuAlaAsnLysAlaGluIleProAlaGluSer 694  
QY 1723 -----ATTTCGATCGATCGGAGGGAAA 1746  
Db  
QY 695 GlnAlaLysGluAsnSerLysProThrCysAsnSerIleHisLeuGlySerGlyAlaLys 714  
QY 1747 ATAACGACATTAAGAGCTAAAGCAGGCGATCAGATCTCTTTAAATGATCCCATC----- 1800  
Db  
QY 715 IleSerGlnLeuArgAlaGlnThrGlyGlnThrIlePhePheThrAspProIleThrThr 734  
QY 1801 -----GAGATGGCAAC 1812  
Db  
QY 735 ThrAlaProAlaAlaAlaValThrAlaLysGlnProGluAlaSerLeuAlaLys 754  
QY 1813 CGAAATAAC-----CAGCAGCGCAGTCT 1836  
Db  
QY 755 AlaThrSerGlyIleProAlaSerAlaSerAlaValSerValProAlaProAlaProThr 774  
QY 1837 TCCAAACTTTAAATTAAC----- 1857  
Db  
QY 775 LysThrProLeuLysIleAsnAlaProAspThrGlnAspProGluIleGlnLysValAla 794  
QY 1858 -----GATGGTGAAGGATACACAGGGATATGTTTTGCT----- 1893  
Db  
QY 795 AlaGluAlaAlaGlnGlnSerAlaValTyrAsnGlyLysIleValPheSerGlyGluLys 814  
QY 1894 -----AATGGAAGCAGTACTTTGTACCAAAATGTT 1923  
Db  
QY 815 LeuSerSerGluAspAlaLysAsnProLeuAsnAlaThrSerValIleHisAsnAspVal 834  
QY 1924 ACGATAGACAGCAAGGAGTGTTCGTGTAAGCAAAATATATCAGTGAATTCCTTA 1983  
Db  
QY 835 SerLeuGluAlaGlyThrLeuValLeuSerSerGlyAlaGlyLeuLeuValAspSerPhe 854  
QY 1984 AGTCAGACAGCTGGAGTCTGTATATCGAAGCTGGAGTACATGGGATTT-----GTA 2037  
Db  
QY 855 ThrGlnLysGluGlySerLeuIleValMetAspGlyGlyThrSerIleIleThrAsnVal 874  
QY 2038 ACTCCA-----CAACCACCAACAGCTCTCTGCC----- 2067  
Db  
QY 875 ThrProAlaSerGluGlyLeuGlnSerArgSerThrProProSerProLysAsnAlaIle 894  
QY 2068 -----GCTAATCAGTTG----- 2079  
Db  
QY 895 ProValIleArgAlaValSerLysValIleAlaSerSerLeuIleAsnLeuArgGluArg 914  
QY 2080 ----- 2082  
Db  
QY 915 AlaAspSerGlyAlaGlyAlaValProThrIleGluGluSerProAspGlySerIle 934  
QY 2083 ACGCTTCCAACTCGCATTTGCTCTCTTCTTTGTTAGCAACAATGAGTTACGAAT 2142  
Db  
QY 935 ThrIleThrAsnLeuAlaValAsnLeuAspSerLeuGluAsnGlyLysValIleThr--- 953

2143 CCTCTACCAATCTCCAGCGCAAGATTCTCATCTCGCAGTCATTGTGTAGCACAACTGCT 2202  
Db  
QY 954 -----LeuAlaAlaLysGlySer 960  
QY 2203 GGTCTCTTCAATAGTGGCCCTATCTTTTGGAGATTGGATGATACAGCTTATGAT 2262  
Db  
QY 961 GlySerValThrLeuThrGlyAspGlnPheGlnAspSerSerGlnAsnPheTyrAsp 980  
QY 2263 AGGTATGATGGCTAGGTTCTAATCAAAAATCAATGTCCTGMAATACAGTTAGGAGCT 2322  
Db  
QY 981 Asn---ProLeuLeuAsnLysAsnPheThrLeuAsnPheLeuAspIle-----Ser 996  
QY 2323 AAGCCCCAGCTAAT-----GCCCATCAGATTGAT 2355  
Db  
QY 997 AlaProAspAlaGluLysIleHisThrGluGlyPheAsnIleIleProGlnGlyAlaThr 1016  
QY 2356 CTAGGAATGAGATGCGCTAAGTATGGCTATCAGGAAGCTGGAAGCTTGGCGGATCCT 2415  
Db  
QY 1017 SerSerAsn-----LeuGlyTyrGlnGlyLysTrpGluValThrGluValLys 1032  
QY 2416 AATACAGCAATATGGTCTTATCTCTGAAAGCTACATGAGCTAAACCTGGTATAAT 2475  
Db  
QY 1033 AspSerSerGlyLysValSerPheGluMetLys-----TrpValSerThrGlyTyrIle 1050  
QY 2476 CTGGGCTGAGCGAGTAGTCTTTTGGTTCCAATAGTTTATGGGATCCATTTTATGAT 2535  
Db  
QY 1051 ProThrAlaAsnArgArgAlaThrLeuValProAsnSerValTrpCysSerAlaIleAsp 1070  
QY 2536 ATACAGTCTGGCATTTACGAATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTCGAGA 2595  
Db  
QY 1071 MetArgAlaPheGlnAsnLeuValGluValSerThrGluGlyGluAspPheHisArgIly 1090  
QY 2596 TTATCGGCTTCTGGAGTTTCGAATTTCTATCATGACCGCGATGCTTTAGTCAGGCA 2655  
Db  
QY 1091 LeuTrpIleSerGlyIleSerAsnPheHisLysAspSerThrLysValGlnGluGly 1110  
QY 2656 TATCGGTATATTAGTGGGTTTATCTCTAGGACAAACTCTCTAC-----TTTGGATCATCG 2712  
Db  
QY 1111 PheArgHisIleSerSerGlyTyrValValGlyValSerThrGlnProIleSerAsnLys 1130  
QY 2713 ATGTTTGGTCTAGCATTTACGAAGTATTTGTAGATCTTAAAGATTTATGATGTCGCT 2772  
Db  
QY 1131 ValMetAspLeuAlaPheCysGlnMetLeuGlyLysSerLysAspTyrArgLeuAlaAsp 1150  
QY 2773 TCCATCATCATGCTGTGCATAGATCCGTTTATCTATCTACCAACAAGCTTTATGTGGA 2832  
Db  
QY 1151 AlaArgSerHisValTyrAlaAlaSerIleHisThrLysCysGluLys---LeuValAsn 1169  
QY 2833 TCCTATTTCTTCGGA----- 2847  
Db  
QY 1170 HistThrPheSerLysArgLysGlyAlaIleLeuAlaArgLysProGluLysSerPro 1189  
QY 2848 -----GATGCGTTTATCCGTCTAGCTACGCGTTGGGAATCAGCATATGAAAACC 2898  
Db  
QY 1190 IleIlePheAspAla-----GlnLeuSerTyrSerLeuSerHisAsnSerMetThrThr 1207  
QY 2899 TCATATACATTTGCGAGAGAGCGATGTTGTTGGGATAATACTGTCTGCTGCGAGAG 2958  
Db  
QY 1208 LysHisThrProAsnProSerSerArgGlyLysTrpAsnAsnHisCysValAlaGlyGlu 1227  
QY 2959 ATTGGAGCGGATTCAGGATTTGATTACT---CCATCTAAGCTCTATTGTAATGATTG 3015  
Db  
QY 1228 LeuGlySerTyrLeuProIleLeuValAspHisProIle-----IleGluGluLeu 1244  
QY 3016 CGTCTCTTCGTCAGCTGAGTTTCTTATGCGGATCATGAATCTTTT-----ACAGAG 3069  
Db  
QY 1245 PheProPheValLysLeuHisIleValPheValGlnGlnGluAspPheLysGluThrGln 1264  
QY 3070 GAAGCGCATCAAGCTCGGCAATTCAGAGCGGACATCTCTCTAATCTATCAGTCTCTGTT 3129  
Db  
QY 1265 GlyGlyThrGluAsnArgAsnPheGlnSerAlaHisPheValAsnValSerLeuProLeu 1284

```

QY 3130 GGAGTGAAGTTTGATCGATGTTCTAGTACACATCCCTAATAATAGCTTTTATGCGGCT 3189
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
1295 GlyValArgPheGluYserThrAsnLysLeu-----AenThrYrAsnIleArgLeuAla 1302
QY 3190 TATATCTGTGATGCTTATCGACCATCTCTGTGATCGACAGCAAGCTCTCTATCCCATCAA 3249
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
1303 TyrGlnProAspIleTyrArgAspAlaProLysSerLysValPheLeuProSerValHis 1322
QY 3250 GAGCATGACACACAGATGCCCTTTCATTTAGACAGACAGATGGTGTGGTAGAGGATCT 3309
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
1323 ThrAlaTrpSerThrGlyAlaThrAsnLeuSerArgGlnAlaMetIleLeuAspGlySer 1342
QY 3310 ATGATGCTCTCTTAACAGTAAGTAATATAGATATATGCCATGAGATATAGATATCGA 3369
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
1343 AspHisHisLeuThrAspAsnLeuGluValPheCysHisGlyAlaPheGluLeuArg 1362
QY 3370 GATGCTTCTGAGGCTGATGGTTGATGCGACAGTAGATCCGGTTC 3417
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
1363 GlySerSerArgAsnTyrAsnValAspIleGlyAlaGlyTyrLysPhe 1378

RESULT 26
PMP7_CHLPPN
ID PMP7_CHLPPN STANDARD; PRT; 936 AA.
AC Q92898; Q9J542; Q9Z5Q4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmp7 precursor (Polymorphic membrane
protein 7) (Outer membrane protein 12).
GN Name=pmp7; Synonyms=omp12; OrderedLocusNames=CP0445, CP0308, CP0462;
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
EX MEDLINE=99206506; PubMed=10192388; DOI=10.1038/7716;
RA Kaiman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
OI Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
EX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Uterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.P., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwin M.L., Nelson W.C., DeBoy R.T., Kolonay J.P.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
EX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
EX Geng M.M., Schuhmacher A., Muehlendorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE OF 658-936.
RC STRAIN=CWL029 / VR1110;

```

```

RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
(Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -1- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
EMBL; AB001627; AAD18589.1; -; Genomic DNA.
EMBL; AB002193; AAF38165.1; -; Genomic DNA.
EMBL; BA000008; BAA98653.1; -; Genomic DNA.
EMBL; AB017158; AAF98393.1; -; Genomic DNA.
EMBL; AJ133034; CAB37067.1; -; Genomic DNA.
DR PIR; B81591; B81591.
DR PIR; C72078; C72078.
DR PIR; C86546; C86546.
DR PHCI-2DPAGE; Q92898; -.
DR TIGR; CP0308; -.
DR InterPro; IPR005546; Auto transpbeta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam_PMP; 7.
DR Pfam; PF07548; ChlamPMP M; 1.
DR TIGRFAMs; TIGR01414; autotrans bar1; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 5.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 23 Potential.
FT CHAIN 24 936 Probable outer membrane protein pmp7.
FT CONFLICT 658 666 PTRHGPHI -> EDNIRYRN (in Ref. 5).
FT CONFLICT 822 822 Y -> H (in Ref. 1, 4 and 5).
SQ SEQUENCE 936 AA; 100107 MW; 3981DB3C950AF95A CRC64;

Alignment Scores:
Pred. No.: 3e-60 Length: 936
Score: 1040.00 Matches: 310
Percent Similarity: 44.1% Conservative: 162
Best Local Similarity: 29.0% Mismatches: 406
Query Match: 13.2% Indels: 192
DB: 1 Gaps: 31

US-10-701-844-1 (1-4435) x PMP7_CHLPPN (1-936)
QY 382 ATGCAACGCTTTCCATAGTCTTCTTTCATGATCTTAGCTATTCTTCTGCTCT 441
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
1 MetLysSerValSerTrpLeuPhePheSerIleProLeuPheSerSerLeuSer 20
QY 442 TTAATGGGGGGGATATGCGACAGCAAAATCATGTT-----CCTCAAGGAATTTAGCAT 495
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
21 Ile-----ValAlaAlaGluValThrLeuAspSerSerAsnSerTyrAsp 36
QY 496 GGGAGAGCTTAACCTGATCATTTCCCTATCTCTTATAGAGATCCGAGTGGGACTACT 555
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
37 GlySerAsnGlyThrPheThrValPheSerThrThrAspAlaAlaGlyThrThr 56
QY 556 GTTTTCTTCGAGGAGAGTTAAACATTAATAATCTTGCAATCTTATTCGAGCTTTCCT 615
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
57 TyrSerLeuLeuSerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAla 76
QY 616 TTAGATTGTTTGGGAACCTATTAGGGAGGTTTCTGTTTATAGGAGGAGGACACTCGTTG 675
Db |||||:|||||:|||||: ||| |||||: ||| |||||: ||| |||||: ||| |||||: |||
77 SerGlyCysPheLeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeu 96

```





Db 690 GlnLeuPheAlaArgAspArgAsnHisIleThrGlyLyAsnHisGlyAspThrTyGly 709

Qy 2794 GGATCCGTTTATCTATCTACCAACAAGCTTTA-----TGTGATCTCTATTGTTGCGA 2847

Db 710 AlaSerLeuTyPheHisThrGluGlyLeuPheAlaAsnHisLeuTyrGly 729

Qy 2848 GATCGTTT----- 2856

Db 730 LysAlaThrArgAlaProTyrValLeuSerGluLeuSerGlnIleIleProLeuSerPhe 749

Qy 2857 ---ATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAACCTCATATACATTGCA 2913

Db 750 AspAlaLysPheSerTyLeuHisThrAspAsnHisMetLysThrTyThr----- 767

Qy 2914 GAGGAGCGGATGTCGT-----TGGATTAATACTCTCGGCTGGGAGAGATTGGACG 2967

Db 768 AspAsnSerIleIleLysGlySerTyrArgAsnAlaPheCysAlaAspLeuGlyAla 787

Qy 2968 GGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGTAATGAGTTGCGTCTTCGTG 3027

Db 788 SerLeuProPheValIleSerValProTyrLeu---LeuLysGluValGluProPheVal 806

Qy 3028 CAAGCTGAGTTTCTTATCGCATCATGAATCTTTTACAGAGGAGCGCATCAAGCTCGG 3087

Db 807 LysValGlnTyrIleTyAlaHisGlnGlnAspPheTyGluArgTyAlaGluGlyArg 826

Qy 3088 GCATTCAAGAGCGGACATCTCTAAATCATCATGTTCTCTGTTGGAGTGAAGTTTGATGCA 3147

Db 827 AlaPheAsnLysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArg 846

Qy 3148 TGTTCTAGTACATCTCTAATAATAGCTTTATGGCGGCTTATATCTGATGCTGCTAT 3207

Db 847 AspSerLysSerGluLysGlyThrTyAspLeuThrLeuMetTyIleLeuAspAlaTyr 866

Qy 3208 CGCACCATCTCTGTTACTGAGACACGCTCTATCCATCAAGACATGACACACAGAT 3267

Db 867 ArgArgAsnProLysCysGlnThrSerLeuIleAlaSerAspAlaAsnTyrMetAlaTyr 886

Qy 3268 GCCTTTCATTAGCAAGCATGAGTGTGTGTAGAGATCTATGATGCTTCTCTAACA 3327

Db 887 GlyThrAsnLeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsn 906

Qy 3328 AGTAATAGATATATGCGCATCGGATGATGATGATGATGATGATGATGATGATGATGAT 3387

Db 907 ProHisMetGluIlePheGlyGlnPheAlaPheGluValArgSerSerArgAsnTyr 926

Qy 3388 GGTGTGAGTCAGGAGTAGAGTCGGTTC 3417

Db 927 AsnThrAsnLeuGlySerLysPheCysPhe 936

RESULT 27

PMP9 CHLPN

ID PMP9 CHLPN STANDARD; PRT; 928 AA.

AC Q92398;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Probable outer membrane protein pmp9 precursor (Polymorphic membrane protein 9) (Outer membrane protein 10)

OS Name=pmp9; Synonyms=omp10; OrderedLocusNames=CP0447, CP0306, CPB0464;

GN Chlamydia pneumoniae (Chlamydophila pneumoniae).

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.

OX NCBI\_TaxID=83558;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=CWL029 / VR1310;

EX MEDLINE=2007584; PubMed=10539856;

RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P.,

RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;

RT "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity."

RL Am. Heart J. 138:8491-8495(1999).

EN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=CWL029;

RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;

RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,

RA Olinger L., Greenwood J., Davis R.W., Stephens R.S.;

RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";

RL Nat. Genet. 21:385-389(1999).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=AR39;

RX MEDLINE=2010255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,

RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Barry K.J.,

RA Bass S., Linher K.D., Weidman J.P., Khouri H.W., Craven B., Bowman C.,

RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,

RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;

RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

RN [4]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=J138;

RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,

RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;

RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";

RL Nucleic Acids Res. 28:2311-2314(2000).

RN [5]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=TW-183;

RX Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,

RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;

RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).

CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.

CC -1- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; AJ133034; CAB37069.1; -; Genomic DNA.

CC EMBL; AE001628; AD18591.1; -; Genomic DNA.

CC EMBL; AE002392; AAF38163.1; -; Genomic DNA.

CC EMBL; BA000008; BAA98655.1; -; Genomic DNA.

CC EMBL; AE017158; AAP98395.1; -; Genomic DNA.

CC PIR; B72077; B72077.

CC PIR; B86546; B86546.

CC TIGR; CP0306; -.

CC InterPro; IPR005546; Auto\_transptbeta.

CC InterPro; IPR006315; Autotransporter.

CC InterPro; IPR011427; ChlamPMP\_M.

CC InterPro; IPR003368; Chlamydia\_PMP.

CC Pfam; PF03797; Autotransporter; 1.

CC Pfam; PF02415; Chlam\_PMP; 7.

CC Pfam; PF07548; ChlamPMP\_M; 1.

CC SMART; SM00710; Pbh1; 5.

CC TIGRFAMs; TIGR01414; autotrans\_bar1; 1.

CC TIGRFAMs; TIGR01376; POMP\_repeat; 6.

KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.

FT SIGNAL 1 26 Potential.

FT CHAIN 27 928 Probable outer membrane protein pmp9.

SQ SEQUENCE 928 AA; 98333 MW; 58910A8F04F12219 CRC64;

Alignment Scores: 5.46e-59 Length: 928

Pred. No.:

Score:	1021.00	Matches:	313
Percent Similarity:	44.6%	Conservative:	161
Best Local Similarity:	29.5%	Mismatches:	404
Query Match:	13.0%	Indels:	184
DB:	1	Gaps:	39

US-10-701-844-1 (1-4435) x PMP9\_CHLPN (1-928)

QY	382	ATGCAACGCTCTTTCATAGTCTCTTTCTTCAATGATCTAGCTTATCTTGCTGCTCT	441
Db	1	MetLysSerSerLeuHisTrpPheLeuIleSerSerLeuAlaLeuPro---LeuSer	19
QY	442	TTAAATGGGGGGATATCAGCAGAAATCATGGTT-----CCTCAAGGAATTTAC	492
Db	20	LeuAsnPheSerAlaPheAlaValValGluIleAsnLeuGlyProThrAsnSerPhe	39
QY	493	GATGGGGAG---ACGTTAACTGTATCATTTCCCTATPACTGTTATAGAGATCCGATGGG	549
Db	40	SerGlyProGlyThrTyrThr-----ProAlaGlnThrThrAsnAlaAspGly	56
QY	550	ACTACTGTTTTTCTGCAGAGAGTTAACTAAATAAAATCTTGACAAATTTCTATCCAGCT	609
Db	57	ThrIleTyrAsnLeuThrGlyAspValSerIleThrAsnAla---GlySerProThrAla	75
QY	610	TTGCCCTTAAAGTTGTTTGGAACTTATTTAGGAGTTTACTGTTTTCAGGAGGACAC	669
Db	76	LeuThrAlaSerCysPheLysGluThrThrGlyAsnLeuSerPheGlnGlyHisGlyTyr	95
QY	670	TCGTTGACTTTCAGAAACATACGGACTTCTACAAATGGGGCAGCTCTAAAGTAATACGCT	729
Db	96	GlnPheLeuLeuGlnAsnIleAspAlaGlyAlaAsn---CysThrPheThrAsnThrAla	114
QY	730	GCTGATGGACTGTTTACTATGAGGTTTAAAGAAATATCTCTTTTCCAAATTCGAATTC	789
Db	115	AlaAsnLysLeuSerPheSerGlyPheSerTyrLeuSerLeu-----	129
QY	790	TTACTTGGCGTACTGCTGTCACACGACTTAATAAGGGTAGCCAGACTCCGACGACAA	849
Db	130	-----IleGlnThrThrAsn	134
QY	850	TCTACACCGTCTAATGCTACTATTTATTTCTAAACACAGATCTTTTGTACTCAATAATAG	909
Db	135	AlaThrThrGlyThrGlyAlaIleLysSerThrGlyAlaCysSerIleGlnSerAsnTyr	154
QY	910	AGTTCTCATTTCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAC	969
Db	155	SerCysTyrPheGlyGlnAsnPheSerAsnAspAsnGlyAlaLeuGlnGlySerSer	174
QY	970	TTAACGGTTCAAGGATTACCAAGCTTTGTTGTTCTTCCAAAGAAATACGCTCAAGCTGAT	1029
Db	175	IleSerLeu---SerLeuAsnProAsnLeuThrPheAlaLysAsnLysAlaThrGlnLys	193
QY	1030	GGGGAGCTTGTCAAGTAGTCACAGTGTCTCTGCTATGCTAGCTAACAGGCTCTATTGCC	1089
Db	194	GlyGlyAlaLeu-----TyrSerThr	200
QY	1090	TTTGTAGCGAATGTTGACGAGTAGAAGGGGGAGGATGCTGCTGTTCCAGATGGGCG	1149
Db	201	-----GlyGlyIleThr	204
QY	1150	CAGGAGTGTCAATCATCTACTTCAACAGAGATCCAGTAGTAACTTTTCCAGAAATACT	1209
Db	205	-----IleAsnAsnThrLeuAsnSer-----AlaSerPheSerGluAsnThr	218
QY	1210	CGGTAGAGTTTGTAGGAACTGATCCGAGTAGGAGGAGTTTACTCTACGCGGAAC	1269
Db	219	Ala-----AlaAsnAsnGlyGlyAlaIleTyrThrGluAlaSer	231
QY	1270	GTTGCTTTCTGTAATAATGGAACCTTG---TTTCTCAACAATGTTGCTTCTCTGTT	1326
Db	232	---SerPheIleSerSerAsnLysAlaIleSerPheIleAsnAsn-----	245
QY	1327	TACATTGCTGTAAACCAACCAAGTGGACAGGCTTCTAATACGAGTAATAATTACGGA	1386

Db	246	SerValThrAlaThrSerAlaThr-----	253
QY	1387	GATGGAGAGCTATCTTCTGTAAAGATCGTGGCAAGCAGGATCCAATAACTCTCGATCA	1446
Db	254	---GlyGlyAlaIleTyrCys-----SerSerThrSerAlaProLysProVal	268
QY	1447	GTTTCCTTTGATGGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGTCTGGGAAAGGG	1506
Db	269	LeuThrLeuSerAspAsnGlyGluLeuAsnPheIleGlyAsnThrAlaIleThrSerGly	288
QY	1507	GGAGCTATTTATGCCAAAAGCTCTCGGTTGCTTAAGTGGCCCTGTACAAATTTTAAGG	1566
Db	289	GlyAlaIleTyrThrAspAsnLeuValLeuSerSerGlyProThrLeuPheLysAsn	308
QY	1567	AATATCGCTTAATGAT-----GGTGAGCGATTTATTTAGGAGAACTCTGA	1611
Db	309	AsnSerAlaIleAspThrAlaAlaProLeuGlyGlyAlaIleAlaIleAlaAspSerGly	328
QY	1612	GAGCTCAGTTATCTGCTGATTATGAGATATTTATTTTCGATGGGAATCTTAAAGAAC	1671
Db	329	SerLeuSerLeuSerAlaLeuGlyAspIleThrPheGluGlyAsn-----Thr	345
QY	1672	GCCAAAGAGAAATGCTGCCGATGTTAATGCGGTAACTGTGCTCTCACAGCCATTTGATG	1731
Db	346	ValValLysGlyAlaSer-----SerSerGlnThrThrThrArgAsnSerIleAsnIle	363
QY	1732	GGA---TCGGGAGGAAATAACGACATTAAGAGCTTAAAGCAGGCGATCAGATTTCTTT	1788
Db	364	GlyAsnThrAsnAlaLysIleValGlnLeuArgAlaSerGlnGlyAsnThrIleTyrPhe	383
QY	1789	AATGATCCATCCAGATGGCAACGGAATAACACGACCGACGCGATCTTCCAACTCTTA	1848
Db	384	TyrAspProIleThr-----ThrSerIleThrAlaAlaLeuSerAspAlaLeu	399
QY	1849	AAAAATTAACGATGGTGAA-----GGATACACAGGGAGATATGTTTGTGCT	1893
Db	400	AsnLeuAsnGlyProAspLeuAlaGlyAsnProAlaTyrGlnGlyThrIleValPheSer	419
QY	1894	-----AATGGAAGCAGTCTTTGTAC	1914
Db	420	GlyGluLysLeuSerGluAlaGluAlaAlaAspAsnLeuLysSerThrIleGln	439
QY	1915	CAAAATGTTACGATAGCAAGGAGGATTTCTTCGTCGAAAAGGCAAAATATCAGTG	1974
Db	440	GlnProLeuThrLeuAlaGlyGlyGlnLeuSerLeuLysSerGlyValThrLeuValAla	459
QY	1975	AATTTCTTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGGGAGTACATGGAT	2031
Db	460	LysSerPheSerGlnSerProGlySerThrLeuLeuMetAspAlaGlyThrThrLeuGlu	479
QY	2032	TTTGTAACTCCACACACCAACCAACGCTCTCCGCGCTAATCAGTTGATCAGCTTTCC	2091
Db	480	ThrAlaAspGly-----IleThrIleAsn	487
QY	2092	AATCTCATTTGCTCTTTCTTTCTTTAGCAACAATGCAGTTACGAATCCTCCTACC	2151
Db	488	AsnLeuValLeuAsnValAspSerLeu-----	496
QY	2152	AATCCTCCAGCGCAAGATTTCTCATCTCGCAGTCATTTGGTAGCAACAATGCTGGT	2208
Db	497	-----LysGluThrLysLysAlaThrLeuLysAlaThrGlnAlaSerGlnThr	512
QY	2209	GTTCAATTTAGTGGGCTATCTTTTTTGGAGTATTCGATGATACAGCTTATGATGAT	2268
Db	513	ValThrLeuSerGlySerLeuSerLeuValAspProSerGlyAsnValTyrGlyAspVal	532
QY	2269	GATTTGGCTAGTTCTTAATCAAAAATCAATGTCTCGAATTTACAGTTAGGAGCTAAGCCC	2328
Db	533	SerTrp-----AsnAsnProGlnVal---PheSerCysLeuThrLeuThrAlaAspAsp	549
QY	2329	CCAGCTAATGCCCA---TCAGATTGACT-----CTAGGGAATGATGCTTAAG	2376





Db 580 LeuSerPheIleAenAenThrAlaLysArgSerGlyGlyGlyIleTyrAlaProLysCys 599  
QY 1453 ---TTTGATGAGAGGGAGTAGTTTCTTTTAGTAGCAATGATGCTGGGAAAGGGGA 1509  
Db 600 ValIleSerGlySerGluSerIleAenPheAepGlyAenThrAlaGluThrSerGlyGly 619  
QY 1510 GCTATTTATGCCAAAAGCTCTCGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAT 1569  
Db 620 AlaIleTyrSerLysAenLysSerIleThrAlaAenGlyProValSerPheThrAenAen 639  
QY 1570 ATCGTAATGATGGTGAGCGCATTTATTTAGGAGAACTCGAGAGCTCAGTTTATCTGCT 1629  
Db 640 SerGlyGlyLysGlyGlyAlaIleTyrIleAlaAepSerGlyGluSerLeuGluAla 659  
QY 1630 GATTATGAGATATTTATTTTCGATGGGAATCTTAAAGAACACGCCAAGAGAAATGCTGCC 1689  
Db 660 IleAepGlyAepIleThrPheSerGlyAen-----ArgAlaThrGluGlyThrSer 676  
QY 1690 GATGTTAATGGCTAACTGTGCTCCACAGCCATTTGATGGGATCGGAGCGGAAATA 1749  
Db 677 ThrProAen-----SerIleHisLeuGlyAlaGlyIle 689  
QY 1750 ACGACATTAAAGAGCTAAAGCGGCATCAGATTCTCTTTAATGATCCCATCGAGATG--- 1806  
Db 690 ThrLysLeuAlaAlaAepProGlyHisThrIleTyrPheTyrAepProIleThrMetGlu 709  
QY 1807 -----GCAACGGG----- 1815  
Db 710 AlaProAlaSerGlyGlyThrIleGluGluLeuValIleAenProValValLysAlaIle 729  
QY 1816 -----AATAACGACGCGGAGTCTTCAAACTCTTAAATTAAC 1857  
Db 730 ValProProGlnProLysAenGlyProIleAlaSerValProValValProValAla 749  
QY 1858 GATGTGAAGGATACACAGGGGATATCTTTT----- 1890  
Db 750 ProAlaAenProAenThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAep 769  
QY 1891 -----GCTAATGAAGACGACTTTGTACCAAAATGTTACGATAGACAGCA 1938  
Db 770 AlaSerIleProAlaAenThrThrThrIleLeuAenGlnLysIleAenLeuAlaGlyGly 789  
QY 1939 AGGATGTTCTTCGTAAAGCGCAAAATATCATGGAATCTCTAAGTCACAGAGGTGG 1998  
Db 790 AenValValLeuLysGluGlyAlaThrLeuGlnValTyrSerPheThrGlnProAep 809  
QY 1999 AGT---CTGTATATGAAGCTGGGAGTACATGGGATTTGTAACTCCACACACCACAA 2055  
Db 810 SerThrValPheMetAepAlaGlyThrThrLeuGluThrThrThrAen----- 826  
QY 2056 CAGCTCTCGCGCTAATCAGTTGATCAGCTTTTCCAACTGCAATTTGTCTCTTTCTCT 2115  
Db 827 -----AenThrAepGlySerIleAepLysAenLeuSerValAenLeuAepAla 843  
QY 2116 TTGTAGCAAAATGAGTTACGATCCTCTCTACCAATCTCCAGCCCAAGATTTCTCAT 2175  
Db 844 LeuAepGlyLysArgMetIleThr----- 851  
QY 2176 CTGCAGTCATTGGTAGCACAACCTGCTGTTTCAATTTAGTGGGCTATCTTTT 2235  
Db 852 -----IleAlaValAenSerThrSerGlyGlyLeuLysIleSerGlyAepLeuLysPhe 869  
QY 2236 GAGGATTTGGATGATACGCTTATGATAGGTATGATGGTCTTAATCAAAAATC 2295  
Db 870 HisAenAenGluGlySerPheTyrAepAen-----ProGlyLeuLysAlaAenLeu 886  
QY 2296 AATGCTCTGAATATACAGTTA---GGGACTAAGCCCCCAGCTAAT----- 2337  
Db 887 AenLeuProPheLeuAepLysSerSerThrSerGlyThrValAenLeuAepPheAen 906  
QY 2338 ---GCCCATCAGATTCTTCACTCTAGGGAATGAGATGCTTAAGTATGGCTATCAGGAGC 2394  
Db 907 ProIleProSerSerMetAla-----AlaProAepTyrGlyThrGlnGlySer 922

QY 2395 TGGAAAGCTTGGTGGGATCCTTAATACAGCAAAATTAATGGTCTTATATCTCTGAAAGCTACA 2454  
Db 923 TrpThrLeuVal-----ProLysValGlyAlaGlyLysValThrLeuValAlaGlu 940  
QY 2455 TGGACTAAACTGGGTATATCCCTGGGCTGAGCGAGTAGCTTCTTTGGTTCCTCAATAGT 2514  
Db 941 TrpGlnAlaLeuGlyTyrThrProLysProGluLeuArgAlaThrLeuValProAenSer 960  
QY 2515 TTATGGGATCCATTTTAGATATACGATCTGGCATTCAGCAATTCAGCAAGTGTGGAT 2574  
Db 961 LeuTrpAenAlaTyrValAenIleHisSerIleGlnGluIleAlaThrAlaMetSer 980  
QY 2575 GGGCGCTCTTATTCGAGGATTAATGGTTTCTCGAGTTTCGAATTTCTTCTATCATGAC 2634  
Db 981 AspAlaProSerHisProGlyIleTrpIleGlyIleGlyAenAlaPheHisGlnAep 1000  
QY 2635 CGCGATGCTTTAGTACGGATATCGGTATATTAGTGGGGTTATCTCTTAGGAGCAAC 2694  
Db 1001 LysGlnLysGluAenAlaGlyPheArgLeuIleSerArgGlyTyrIleValGlyGlySer 1020  
QY 2695 TCCTACTTTGGATCATCG---ATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGACT 2751  
Db 1021 MetThrThrProGlnGluTyrThrPheAlaValAlaPheSerGlnLeuPheGlyLysSer 1040  
QY 2752 AAAGATTATGATGTGCTGCTTCCAATCATCATGCTTGCATAGGATCCGTTTATCTATCT 2811  
Db 1041 LysAepTyrValValSerAepIleLysSerGlnValTyrAlaGlySer-----LeuCys 1058  
QY 2812 ACCCAACAGCTTTATGATGCTATTTGTTGGAGATGCGTTTATCCGTGCT----- 2865  
Db 1059 AlaGlnSerSer-TyrValIleProLeuHisSerSerLeuArgHisValLeuSerLys 1078  
QY 2866 -----AGCTACGG 2873  
Db 1078 sValLeuProGluLeuProGlyGluThrProLeuValLeuHisGlyGlnValSerTyrGln 1098  
QY 2874 GTTTGGGAATCAGCATATGAAACCTCATATACATTTGCAGAGGAGCGGATGTTGTTG 2933  
Db 1098 YArgAenHisHisAenMetThrThrLysLeuAlaAenAenThrGlnGlyLysSerAepTyr 1118  
QY 2934 GGATAAATACGTCTGCTGGAGATTTGGAGCGGATTTACCGATTTGATTTCTCATTC 2993  
Db 1118 pAepSerHisSerPheAlaValGluValGlyLysSerLeuProValAepLeu-----As 1136  
QY 2994 TAAGCTCTATTGAATGATGCTGCTTCTGTCGAAGCTGAGTTTCTTATGCGCATCA 3053  
Db 1136 nTyrArgTyrLeuThrSerTyrSerProTyrValLysLeuGlnValValSerValAenGln 1156  
QY 3054 TGAATCTTTTACAGAGAGCGGATCAAGCTCGGCGATTTCAAGAGCGGACATCTCTAAA 3113  
Db 1156 nLysGlyPheGlnGluValAlaAlaAepProArgIlePheAepAlaSerHisLeuValAs 1176  
QY 3114 TCTATCAGTTCTGTTGGAGTGAAGTTTGTATGATGTTCTAGTACACATCTTAATAATA 3173  
Db 1176 nValSerIleProMetGlyLeuThrPheLysHisGluSerAlaLysProSerAlaLe 1196  
QY 3174 TAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTGAGACAAC 3233  
Db 1196 uLeuLeuThrLeuGlyTyrAlaValAlaAepAlaTyrArgAepHisProHisCysLeuThrSe 1216  
QY 3234 GCTCTTATCCCATCAAGAGACATGACACAGATGCTCTTTTATTAGCAAGATGGAGT 3293  
Db 1216 r---LeuThrAenGlyThrSerTyrPheAlaThrAenLeuSerArgGlnAlaPhe 1235  
QY 3294 TGTGTTAGGAGTCTATGCTTCTTCAACAGTAATAATAGAAGTATATAGCCATGG 3353  
Db 1235 ePheAlaGluAlaSerGlyHisLeuLysLeuHisGlyLeuAepCysPheAlaSerGln 1255  
QY 3354 AAGATATGATGATGCTCTCGAGGCTATGTTTGTAGTCAGAGTAGAGTCCG 3413  
Db 1255 ySerCysGluLeuArgSerSerSerArgSerTyrAenAlaAenCysGlyThrArgTyrSe 1275







RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FW-183;  
 RA Geng M.M., Schuhmacher A., Muehlendorfer I., Bensch K.W., Schaefer K.P.,  
 RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.,  
 RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with  
 RT other Chlamydia strains based on whole genome sequence analysis.";  
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BA000008; BAA98226.1; -; Genomic DNA.  
 DR EMBL; AE017157; AAP97951.1; -; Genomic DNA.  
 DR PIR; H86492; H86492.  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR InterPro; IPR006315; Autotransporter.  
 DR InterPro; IPR005546; Auto transportbeta.  
 DR InterPro; IPR011427; Chlam\_PMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF07548; Chlam\_PMP\_M; 1.  
 DR Pfam; PF02415; Chlam\_PMP; 6.  
 DR TIGRFAMs; TIGR01414; autotrans\_bar1; 1.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 6.  
 SQ SEQUENCE 772 AA; 82931 MW; 484FC56D635801EB CRC64;

## Alignment Scores:

Pred. No.: 2,96e-56 Length: 772  
 Score: 979.50 Matches: 261  
 Percent Similarity: 46.2% Conservative: 144  
 Best Local Similarity: 29.8% Mismatches: 323  
 Query Match: 12.4% Indels: 149  
 DB: 2 Gaps: 24

US-10-701-844-1 (1-4435) x Q9RB71\_CHLPN (1-772)

QY 919 TTCTATAGTAATTTAGTCTCTGGAGTGGGGAGCTATAGATCTTAAGAGCTTAACGGTT 978  
 DB 3 PheSerLysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaIleThrLeuSerLeu 22  
 QY 979 CAAGGAATTAGCAAGCTTGTGTCCTCCAGAAATACTGCTCAAGCTGATGGGGAGCT 1038  
 DB 23 ThrGlyThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGlyAla 42  
 QY 1039 TGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAAACGAGGCTCTATTGCTTTGTAGCG 1098  
 DB 43 IleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyGluValSerPhe 60  
 QY 1099 AATGTTGCAGGAGTAGAGGGAGGAGTTGCTGCTGTTCCAGGATGGGACGAGGAGTG 1158  
 DB 60 ----- 60  
 QY 1159 TCATCATCTACTTCAACAGAGATCCAGTAGTAGTAAATTTTCCAGAAATATCGCGTAGAG 1218  
 DB 61 SerAspAsnThrSerSerAspSer----- 68  
 QY 1219 TTTGATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTTC 1278  
 DB 69 -----GlyAlaAlaIlePheThrGluAlaSerValThrIle 80  
 QY 1279 CTGAATATGGAAACCTTTGTTTCTCAACAATTTGCTTCTCTCTGTTTACATTTGCTGCT 1338  
 DB 81 SerAsnAsnAlaLysValSerPheIleAspAsn----- 91  
 QY 1339 AAGCAACCAACAGTGGACAGGCTTCTTAATACGAGTAGTAATATTACGAGATGGAGGAGCT 1398  
 DB 92 -----LysValThrGlyAlaSerSerSerThrThrGlyAspMet-----SerGlyGlyAla 108  
 QY 1399 ATCTTCTGTAAGATGTGGCCAGCAGGATCCATAACTCTGGATCAGTTTCTCTTTGAT 1458  
 DB 109 IleCysAlaTyrlLys-----ThrSerThrAspThrLysValThrLeuThr 123  
 QY 1459 GGAGAGGAGTAGTATTTCTTTAGTGAATAGTCTGCTGGAAAGGGGAGGAGCTATTAT 1518  
 DB 124 GlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrThrAlaGlyGlyAlaIleTy 143

QY 1519 GCCAAAAAGCTCTCGTTGCTAATGTGGCCTGTACAAATTTTAAAGAAATATCCCTAAT 1578  
 DB 144 VallysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsnSerValaAsn 163  
 QY 1579 -----GATGCTGGAGCGATTTATTTAGGAGATCTGGAGAGCTCGATTTA 1623  
 DB 164 GlyGlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGluLeuSerLeu 183  
 QY 1624 TCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAGAT 1683  
 DB 184 SerAlaAspSerGlyAspIleValPheLeuGlyAsn-----ThrValThrSerThr 200  
 QY 1684 GCTGCCGATGTAATGGCGTAATCTGTCTCCTCACAGCAATTCATGGGATCGGAGGG 1743  
 DB 201 ThrProGlyThrAsn-----ArgSerSerIleAspLeuGlyThrSerAla 215  
 QY 1744 AAAATAACGACATTAAGAGCTAAACAGCGCATCATGATTTCTTTAAATCATCCATCGAG 1803  
 DB 216 LysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleTyrlPheTyrlAspProIleThr 235  
 QY 1804 ATGCCAAACGGAATAAACCCAGCCAGCGAGTCTTCCAAACTTCTAAAAATTAAACGAGGT 1863  
 DB 236 ThrGlySerSerThrThr-----ValThrAspValLeuLysValaAsnGluThr 251  
 QY 1864 GAAGGA-----TACACAGGGATATTGTTTTGCT----- 1893  
 DB 252 ProAlaAspSerAlaLeuGlnTyrlThrGlyAsnIleIlePheThrGlyGlyLysLeuSer 271  
 QY 1894 -----AATGGAACGAGTACTTTGTACCAAAATGTTACGATA 1929  
 DB 272 GluThrGluAlaAspSerLysAsnLeuThrSerLysLeuLeuGlnProValThrLeu 291  
 QY 1930 GAGCAAGGAAGGATTTCTCTGCTGAAAGCGCAAAATATATCAGTGAATTTCTTAAGCTCAG 1989  
 DB 292 SerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThrGlnAlaPheThrGln 311  
 QY 1990 ACAGGTGGGAGT---CTGTATATGGAAGCTGGAGGTACATGGGATTTTGTAACTCCACAA 2046  
 DB 312 GlnAlaAspSerArgLeuGluMetAspValGlyThrThrLeuGlu----- 326  
 QY 2047 CCACCAACAGCCTCTCTCCGCTAATCAGTTGATCAGCTTTTCCATTCGATTTGTCT 2106  
 DB 327 -----ProAlaAspThrSer-----ThrIleAsnAsnLeuValIleAsn 339  
 QY 2107 CTTTCTCTTTGTTGACAAACAAATGCAGTTACGAATCCTCCTACCAATCCTCCAGCGCA 2166  
 DB 340 IleSerSerIle-----Asp 344  
 QY 2167 GATTCTCATCTGCACTCATTTGGTAGC---ACAACTGCTGGTGTCTGTCAATAGTGGG 2223  
 DB 345 GlyAlaLysLysAlaLysIleGluThrLysAlaThrSerLysAsnLeuThrLeuSerGly 364  
 QY 2224 CTTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGATGTTGGCTAGTTCT 2283  
 DB 365 ThrIleThrLeuLeuAspProThrGlyThrPheTyrlGluAsnHisSer---LeuArgAsn 383  
 QY 2284 AATCAAAAATCAATGCTCTGAAATTTACAGTTA---GGGACTAAAGCCCCCAGCTAATGCC 2340  
 DB 384 ProGlnSerTyrlAspIleLeuGluLeuLysAlaSerGlyThrValThrSerThrAlaVal 403  
 QY 2341 CCATCAGATTGACTCTAGGGAATGAGATGCCCTTAAGTATGGCTATCAAGAAAGCTGG--- 2397  
 DB 404 ThrProAspProIleMetGlyGlyLysPhe---HisTyrlGlyTyrlGlnGlyThrTrpGly 422  
 QY 2398 AAGCTTGGGAGATCTTAATACAGCAATAATATGGTCTTATCTCTGAAGAGCTACATGG 2457  
 DB 423 ProIleValTrpGlyThrGlyAlaSerThrThrAlaThrPhe-----AsnTrp 438  
 QY 2458 ACTAAACTGGGTATTAATCTGGCTGAGCGAGTAGTCTTTTGGTTCCAAATAGTTTA 2517  
 DB 439 ThrLysThrGlyTyrlPleProAsnProGluArgIleGlySerLeuValProAsnSerLeu 458  
 QY 2518 TGGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGG 2577

```

Db      459  TrpAenAlaPheIleAspIleSerSerLeuHisTyrLeuMetGluThrAlaAsnGluGly 478
QY      2578  CGCTCTTATGTCGAGGATATAGGTTCTCGAGTTTCGAATTTCTTCTATCATGACCGC 2637
Db      479  LeuGlnGlyAspAlaPheTrpCysAlaGlyLeuSerAsnPhePheHisLysAspSer 498
QY      2638  GATGCTTTAGGTACGAGTATCGGTATATAGTGGGGTTATCTCTTAGGAGCAATCC 2697
Db      499  ThrLysThrAgaGlyPheArgHisLeuSerGlyGlyTyrValIleGlyGlyAsnLeu 518
QY      2698  TACTTTGGATCA---TCCATGTTTGGTCTAGCATCCGAGTATTCGAGTATTCGAGTCTAAA 2754
Db      519  HisThrCysSerAspLysIleLeuSerAlaAlaPheCysGlnLeuPheGlyArgAspArg 538
QY      2755  GATTATGATGTCGTCGTCCTCAATCATCATGCTTCATGAGTATTCGTTATCTATCTACC 2814
Db      539  AspTyrPheValAlaLysAsnGlnGlyThrValTyrGlyGlyTyrLeuTyr-TyrGlnHi 558
QY      2815  CAACAAGCTTTATGCGATCTATTGTTGTCGAGATGCGTTTATCCGTGCT----- 2865
Db      558  sAsnGluThrTyrIleSerLeuProCysLysLeuArgProCysSerLeuSerTyrValPr 578
QY      2866  -----AGCTACGGGTTTGGGAATCAGCATAT 2891
Db      578  oThrGluIleProValLeuPheSerGlyAsnLeuSerTyrThrHisThrAspAsnAspLe 598
QY      2892  GAAACCTCATATACATTTGCGAGAGGAGCGATGTCGTTGGGATATATACTGTCTGCG 2951
Db      598  uLysThrLysTyrThrTyrProThrValLysGlySerTrpGlyAsnAspSerPheAl 618
QY      2952  TGGAGATTTGGAGCGGATTTACCATGTCGATTAATCTCATCTAAGCTCTATTTGAATGA 3011
Db      618  aLeuGluPheGlyArgAlaProIleCysLeuAspGluSerAlaLeuPhe---GluGI 637
QY      3012  GTTGCGTCTCTTGGCAAGTCAGTTCTTATGCCATCATCATGATCTTTTACAGAGGA 3071
Db      637  nTyrMetProPheMetLysLeuGlnPheValTyrAlaHisGlnGluGlyPheLysGluGI 657
QY      3072  AGGCGATCAAGTCGCGCATCAAGAGCGGACATCTCTTAATCTATCATGTCCTGTTGG 3131
Db      657  nGlyThrGluAlaArgGluPheGlySerSerArgLeuValAsnLeuAlaLeuProIleGI 677
QY      3132  AGTGAAGTTTGCATGCTTCTAGTACACATCTTAATATATAGCTTTATGGCGGCTTA 3191
Db      677  yIleArgPheAspLysGluSerAspCysGlnAspAlaThrTyrAsnLeuThrLeuGlyTy 697
QY      3192  TATCTGTCATGCTTATCCACCATCTCTGCTACTGAGACAACGCTCTCATCCATCAAGA 3251
Db      697  rThrValAspLeuValArgSerAsnProAspCysThrThrThrLeuArgIleSerGlyAs 717
QY      3252  GACATGGACAACAGATGCTTTTCATTATAGCAAGCATGGGTGTGGTTAGAGGATCTAT 3311
Db      717  pSerTrpLysThrPheGlyThrAsnLeuAlaArgGlnAlaLeuValLeuArgAlaGlyAs 737
QY      3312  GTATGCTTCTTAACAAGATATATAGATATATATGCGCATGGAAGATATAGTATCGAGA 3371
Db      737  nHisPheCysPheAsnSerAsnPheGluAlaPheSerGlnPheSerGlnPheSerGluLeuArgGI 757
QY      3372  TGCTTCTCGAGGCTATGTTGTGGCAGGAAGTAGAGTCGGTTC 3417
Db      757  ySerSerArgAsnTyrAsnValAspLeuGlyAlaLysTyrGlnPhe 772

RESULT 31
PMP1 CHLPN
ID -PMP1 CHLPN STANDARD; PRT; 922 AA.
AC Q929G5; Q9K1Y9; Q9Z4H9;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmp1 precursor (Polymorphic membrane
protein 1) (Outer membrane protein 6).
GN Name=pmp1; Synonyms=omp6; OrderedLocusNames=CpN0005, CP0770, CpB0006;

```

```

OS      Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX      NCBI_TaxID=83558;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CWL029 / VR1310;
RX      MEDLINE=20007584; PubMed=10539856;
RA      Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA      Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT      "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT      role in immunopathogenicity.";
RL      Am. Heart J. 138:8491-8495(1999).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=CWL029;
RX      MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA      Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA      Olinger L., Greenwood J., Davis R.W., Stephens R.S.;
RT      "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL      Nat. Genet. 21:385-389(1999).
RN      [3]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=AR39;
RX      MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA      Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
RA      White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA      Bas S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA      Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA      McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT      "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT      pneumoniae AR39.";
RL      Nucleic Acids Res. 28:1397-1406(2000).
RN      [4]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=J138;
RX      MEDLINE=2003049; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA      Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA      Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT      "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT      from Japan and CWL029 from USA.";
RL      Nucleic Acids Res. 28:2311-2314(2000).
RN      [5]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=TW-183;
RX      Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA      Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT      "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT      other Chlamydia strains based on whole genome sequence analysis.";
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC      -1- SURCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC      (Potential).
CC      -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; AJ133035; CAB37075.1; -; Genomic DNA.
CC      EMBL; AB001585; AA018163.1; -; Genomic DNA.
CC      EMBL; AB002237; AAF38570.1; -; Genomic DNA.
CC      EMBL; BA000008; BAA98215.1; -; Genomic DNA.
CC      EMBL; AB017157; AAP97939.1; -; Genomic DNA.
CC      PIR; B72131; B72131.
CC      PIR; B86491; E86491.
CC      TIGR; CP0770; -.
CC      InterPro; IPR005546; Auto_transpbeta.
CC      InterPro; IPR011427; ChlamPMP_M.
CC      InterPro; IPR003368; Chlamydia_PMP.
CC      Pfam; PF03797; Autotransporter; 1.
CC      Pfam; PF02415; Chlam_PMP; 7.
CC      Pfam; PF07548; ChlamPMP_M; 1.

```

DR	TIGR01376;	POMP_repeat; 5.			
KW	Complete proteome; Membrane; Multigene family; Outer membrane; Signal.				
FT	SIGNAL	26	Potential.		
FT	CHAIN	27	922	Probable outer membrane protein pmp1.	
FT	FT	CONFLICT	14	14	F -> L (in Ref. 1).
FT	FT	CONFLICT	251	251	A -> P (in Ref. 5).
FT	FT	CONFLICT	375	375	Y -> C (in Ref. 1).
FT	FT	CONFLICT	606	606	D -> N (in Ref. 1).
FT	FT	CONFLICT	836	836	S -> P (in Ref. 5).
SQ	SEQUENCE	922 AA;	100458 MW;	DFF2AB6333AB031C	CRC64;

  

Alignment Scores:	
Pred. No.:	8 41e-56 Length: 922
Score:	973.00 Matches: 306
Percent Similarity:	41.5% Conservative: 148
Best Local Similarity:	28.0% Mismatches: 253
Query Match:	12.3% Indels: 386
DB:	1 Gaps: 35

  

US-10-701-844-1 (1-4435) x PMP1_CHLPN (1-922)	
Qy	375 TTGTGTGATGCAACCGTCTTCCATAGTCTCTTCTTCTTCAAT-----GAT 419
Db	5 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 21
Qy	420 TCTAGCTATTCTTCTGCTCTTTTAAATGGGGGGGATATGCAGCAGAAATCATGTTGCC 479
Db	22 ThrSerLeu-SerAlaThrThriSer-----LeuThrPr 33
Qy	480 TCAAGGAATTTACGATGGGAGACGTAACTGTATCATTTCCCTATCTGTTATAGGAGA 539
Db	33 cGluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTyAsnVal-----50
Qy	540 TCCGATGGGACTACTGTTTTTCTGCAGGAGATTAACTAAAAAATCTTGACAAATTC 599
Db	51 -GlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 70
Qy	600 TATTCGAGTTTGCCTTTAAGTTGTTTGGGAACTTATTAGGAGGTTTACTGTTTATAG 659
Db	70 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGl 88
Qy	660 GAGAGGACACTGTTGACTTTTCGAGAACATACGGACTTCTACA-----AATGGGGCAGC 713
Db	88 yAsnHisGlyLeuTyPheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVala 108
Qy	714 TCTAAGTAATACGGCTGTGATGAGCTGTTTACTATGAGGGTTTAAAGAATTATCCTT 773
Db	108 lLeuCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128
Qy	774 TTCCAATTGCAATTCAATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 833
Db	128 e-----llegl 130
Qy	834 GACTCCGACGACCAACATCTACACCGTCTAATGGTACTATTTATTCTAAACAGATCTTTT 893
Db	130 nSerPro-----GlyAspIlelySgluGlnGlyCysLeuTySerLysAsnAlaLeuMe 148
Qy	894 GTTACTCAATAATAGGAAGTTCTCATCTATAGTAATATTAGTCTCTGGAGATGGGGGAGC 953
Db	148 tLeuLeuAsnAsnTyValValArgPheGluGlnAsnGlnSerLysThrLysGlyGlyAl 168
Qy	954 TATAGATGCTAAGAGCTTAACGGTTCAAGGAATTTAGCAAGCTTTGTGTCTTCCAGAAAA 101
Db	168 aIleSerGlyAlaAsnValThrIleValGlyAsnTyAspSerValSerPheTyGlnAs 188
Qy	1014 TACTGCTCAAGCTGATGGGGGAGCTTGTCAAGTAGTACCAGGTTTCTCTGCTATGCTAA 1073
Db	188 nAlaAlaThrPheGlyGly-----AlaIleHis 198
Qy	1074 CGAGGCTCTTATTTGCTTTGTACGGAATTTGCGAGGATTAAGAGGGGAGGATTCGTGC 113
Db	198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg-----212

```

Db      487  -----SerSerSerThrAlaAlaValIleLysAlaAsnTh 499
QY      2199 TGCTGGT-----TCTGTACAAATAGTGGCGCTATCTTT-----2232
Db      499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519
QY      2233 -TTTGAGGATTTGGATGATACAGCTTATGATAGGTATGTTGGCTAGGTTCTTAATCAAA 2291
Db      519 eTyrgluAspLeuArg-----MetArgAsnSerGlnTh 530
QY      2292 AATCAATCTCTGAAATATACAGTTAGGACTAAGCCCGCAGCTAATGCCCATCAGATTT 2351
Db      530 rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlyGlySerValThrVa 546
QY      2352 GACTCTAGGGAATGAGATG-----CCTAAGTATGCTATCAAGAACTGGAAGCT 2402
Db      546 lThrAlaGlyAspPheLeuProValSerProHisTyrGlyPheGlnGlyAsnTrpLysLe 566
QY      2403 TGCCTGGGATCCTAATACAGCAATAATGCTCCTTATCTCTGAAGACTACATGGACTAA 2462
Db      566 uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-----PheTrpAspGly 582
QY      2463 AACTGGGTATAATCTCGGCGCTGAGCGAGTAGCTTCTTTGGTCCAAATAGTTTATGGG 2522
Db      582 sIleAsnTyrLysProArgProGluLysGluGlyAsnLeuValProAsnIleLeuTrpGl 602
QY      2523 ATCCATTTTATGATATPACGATCTGCGCAATTCAGCAATTCAGCAAGTGTGGATGGCGTC 2582
Db      602 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuGl 622
QY      2583 TTATGTGCGAGATTTATGGTTTCGGAGTTTCGAAATTTCTTCTATCATGACCGGATGC 2642
Db      622 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPhePheHisValSerAlaSerGl 642
QY      2643 TTTAGGTCAGGATATCGGTATATTAGTGGGGTTTATTCCTTAGGCAAACTCCTACTT 2702
Db      642 uAspAsnIleArgTyrArgHisAsnSerGlyGlyTyrValLeuSerValAsnAsnGluIl 662
QY      2703 TGGATCATCGATGTTT---GGTCTAGCATTTACCGAAGTATTGGTAGATCTAAAGATTA 2759
Db      662 eThrProLysHisTyrThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 682
QY      2760 TGTAGTGTGCTCCATCATCATGCTTGCTAGGATCGTTTATCTATCTATCCCAACA 2819
Db      682 rAlaValSerAsnAsnGluTyrArg-----MetTyrLeu-----693
QY      2820 AGCTTTATGTGATCCTATTGTTTC-----GGAGATGCGTTTATCCG 2861
Db      694 -----GlySerTyrLeuTyrGlnTyrThrThrSerLeuGlyAsnIlePheArgTy 710
QY      2862 TGCTAGC-----2868
Db      710 rAlaSerArgAsnProAsnValAsnValGlyLeuSerArgPheLeuGlnAsnPr 730
QY      2869 -----TACGGTTTCGGATCGATCGCATATGAAAACCTC 2900
Db      730 oLeuMetIlePheHisPheLeuCysAlaTyrGlyHisAlaThrAsnAspMetLysThrAs 750
QY      2901 ATATACATTTGCAGAGGAGCGATGTTCTGGTGGGATAATAACTCTCTGCTGGCTGGAGAT 2960
Db      750 pTyrAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCy 770
QY      2961 TGGAGCGGGATTTACGATTTGATTAATCTCAATGCTTATTTGAATGAGTTGCGTCC 3020
Db      770 sGlyGlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr 790
QY      3021 TTTCTGCGAAGCTGAGTTCTTATTCGCGATCATGATCTTTTACAGAGGAGCGGATCA 3080
Db      790 oPheMetLysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheLysGluThrThrAlaAs 810
QY      3081 AGCTCGGCGATTTCAAGAGCGGACATCTCTCTAATCTATCAGTTCTGTTGGAGTGAAGTT 3140

```

```

Db      810 pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 830
QY      3141 TGATCGATGTTCTAGTACACATCTTAATAATATAGCTTTATGGCGGCTTATATCTGTGA 3200
Db      830 eGluLysLeuAlaLeuSerGlnAspValLeuTyrAspPheSerPheSerTyrIleProAs 850
QY      3201 TGCTTATCGCACCATCTCTGGTACTGAGACAAACGCTCCTATCCCATCAAGAGACATGGAC 3260
Db      850 pIlePheArgLysAspProSerCysGluAlaAlaLeuValIleSerGlyAspSerTrpLe 870
QY      3261 AACAGATGCTCTTTCATTTAGCAACATCGGAGTTGTGTGTAGAGATCTATGTATGCTTC 3320
Db      870 uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyrHi 890
QY      3321 TCTAACAGTAATATAGATATATATGCCCATGGAAGATATGAGATGATCGAGATGCTTCG 3380
Db      890 sPheAsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 910
QY      3381 AGGCTATGGTTGAGTCGACGAGAGTAGAGTCCGGTTC 3417
Db      910 gAsnTyrAsnIleAsnCysGlySerLysPheArgPhe 922

RESULT 32
Q822Q5_CHLCV
ID      Q822Q5_CHLCV PRELIMINARY; PRT; 841 AA.
AC      Q822Q5;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Polymorphic outer membrane protein G family
DE      protein/autotransporter.
GN      OrderedLocustNames=CCA00624;
OS      Chlamydomophila caviae.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
NCBI_TaxID=83557;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=GPIC;
RX      MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA      Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA      Heidelberg J.F., Holtzapfle E.K., Khouri H.M., Federova N.B.,
RA      Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA      White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
RA      Bavoil P.M., Fraser C.M.;
RT      *Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
RT      examining the role of niche-specific genes in the evolution of the
RT      Chlamydiaceae.
RL      Nucleic Acids Res. 31:2134-2147(2003).
DR      EMBL; AB016996; AAP05366.1; -, Genomic_DNA.
DR      TIGR; CCA00624; -.
DR      GO; GO:0019867; C:outer membrane; IEA.
DR      InterPro; IPR006315; AutoTransporter.
DR      InterPro; IPR005546; AutoTranspBeta.
DR      InterPro; IPR011427; ChlamPMP M.
DR      InterPro; IPR003368; Chlamydia_PMP.
DR      Pfam; PF03797; AutoTransporter; 1.
DR      Pfam; PF07548; ChlamPMP M; 1.
DR      Pfam; PF02415; Chlam_PMP; 4.
DR      TIGRFAMs; TIGR01414; autotrans barl; 1.
DR      TIGRFAMs; TIGR01376; POMP_repeat; 3.
KW      Complete proteome.
SQ      SEQUENCE 841 AA; 90411 MW; 21A9ED1680D7C65D CRC64;

Alignment Scores:
Pred. No.: 1.03e-55 Length: 841
Score: 971.50 Matches: 298
Percent Similarity: 40.0% Conservative: 132
Best Local Similarity: 27.7% Mismatches: 347
Query Match: 12.3% Indels: 299
DB: 2 Gaps: 30

US-10-701-844-1 (1-4435) x Q822Q5_CHLCV (1-841)

```

177 -----GlLeuLysPheGluGlyAsnLysTyThrLeuLeuPheSerGly 190  
1486 AATGTAGCTGCTGGGAAGGGGAGCTATTATATGCCAAAAGCTCTCGGTGCTAACTGT 1545  
191 AenSerSerGlnGlnGluGlyAlaIleTyAlaLysLysLeuSerIleSerGly 210  
1546 GGCCCTGCTCAATTTTAAAGGAATATCGCTAAT-----GATCGTGGAGCG 1590  
211 GlyProThrLeuPheSerAsnAsnSerThrSerLysAlaAlaAspProLysGlyGlyAla 230  
1591 ATTTATTTAGGAGAACTCT---GGAGAGCTCAGTTTATCTGCTGATATATGAGATATTATT 1647  
231 IleCysIleAlaAspAlaAspSerGluCysSerLeuThrAlaGluAsnGlyAspIleIle 250  
1648 TTCGATGGGAATCTTAAAGAACAACGCCAAAGAGATGTCGCCGATGTTAATGGCGTAACT 1707  
251 PheAspGlyAsnLysIleIleThrThrGlyThrProSerThrLysArgAsn----- 267  
1708 GTGTCTCTCAACGCCATTCGATGGATCGGAGGAGGAAATAACGACATATAAGAGCTAAA 1767  
268 -----SerIleAspLeuGlySerGlyLysPheSerGlnLeuArgAlaArg 283  
1768 GCAGGCGCATCAGATTCCTTTAATGATCCATCCAGATGGCAACGCGAAATACACAGCCA 1827  
284 AspGlyPheGlyValPhePheTyAspProIle-----AlaAsnAsn----- 297  
1828 GCGCAGCTCTCCAAACTCTTAAAAATTAAACGATGGTGAAGGA-----TACACAGGG 1878  
298 GlySerAspThrAspThrLeuGluIleAsnLysAlaAspGlyAlaAlaThrTySerGly 317  
1879 GATATTGTTTTGCT-----AATGGA 1899  
318 ArgIleValPheSerGlyGluLysLeuThrGluAspGluLysGlnValThrAspAsnLeu 337  
1900 AGCAGTACTCTGTACCAAAATGTTACGATAGACAGGAGGATGTTCTTCCTGGAAGAAG 1959  
338 LysSerPhePheLysGlnProLeuThrValGlySerGlySerPheValLeuLysAsnGly 357  
1960 GCAAAATATCAGTGAATCTCTAAGTCAGACAGGTGGGAGTCTGTATATGGAAGCTGGG 2019  
358 ValThrValSerAlaLysGlnIleThrGlnSerGlyGlyAlaIleGluMetAspAlaGly 377  
2020 AGTACATGGGATTTTGTAATCCACACACCACACAGCTCTCTCGCTAATCAGTTG 2079  
378 ThrAsn----- 379  
2080 ATCAGCGTTTCCAATCTGCATTTGTCTCTTCTTCTTTGTTAGCAACAATGCGATTACG 2139  
380 LeuThrSerThrThrGluAspIleSerLeuSerAsnLeuValIleAsnThrAlaSerLeu 399  
2140 AATCCTCTTACCATCTCCAGGCGAAGATTCATCTCTGCACTCATTTGGTAGCACACACT 2199  
400 GlyGlyGlyGlyValProLeuAlaAlaGln-----IleSerAlaGluGlyThr 415  
2200 GCTGGTCTGTTCATTAATAGTGGCCCTATCTTTTGGAGGATTTGGATGATACAGCTTAT 2259  
416 AsnLysSerValThrIleSerSer-----LeuAsnLeuValAspAlaAspGlyAsnGlyTy 434  
2260 GATAGGTATGATTTGGGTAGTCTTAATCAAAAAATCAATCTCTGAAATTTACAGTTA--- 2316  
435 GluTyProValPheSerThrThrArgGluPheProSerIleIleGluAlaLysAlaAsn 454  
2317 GGGGATTAAGCCCCAGCTAATAGCCCCATCAGATTTGATCTCTAGGGAATGAGATGCCCTAAG 2376  
455 GlyThrGlyThrProThrIleProThrThrHisLeuThrAspHisAlaProAlaAlaHis 474  
2377 TATGGCTATCAAGGAGCTGAAGCTGCTGGGATCCTTAATACAGCAATAATGGTCCT 2436  
475 TyrGlyTyThrGlnGlyLeuThrThrSerTrpAlaGlnGlyThrAlaThrThrSerGln 494  
2437 TATACTCTGAAAGCTACATGCACTAAACCTGGGTATTAATCTGGGCTCTGAGCGAGTAGCT 2496

Db 495 LeuAlaThrLeuAla---TrrpGlnGlnThrGlyTyrAsnProAsnProGluArgGlnGly 513  
 Qy 2497 TCTTTGGTTCCTCAATAGTTATGGGATCCATTTTAGATATACGATCTCGCGATTCAGCA 2556  
 Db 514 ProLeuValProAsnThrLeuTrrpGlySerPheSerAspValArgAlaIleGlnAsnLeu 533  
 Qy 2557 ATTCAAGCAAGTGTGGATGGCGCTCTTATGTCGAGGATATATGGTTCTTCGGAGTTTCG 2616  
 Db 534 MetAspIleSerValAsnGlnAlaAspTyrGlnArgGlyLeuTrrpAlaSerGlyLeuAla 553  
 Qy 2617 AATTTCTTCTATCATGACCGCGATCTTTAGGTCAGGATATCGGTATATAGTGGGGT 2676  
 Db 554 AsnPheLeuGlnLysSerGlyThrGluThrLysArgLysPheArgHisSerAlaGly 573  
 Qy 2677 TATTCCTTAGAGCAAACTCC---TACTTTGATCATGATGTTTGGTCTAGCATTTACC 2733  
 Db 574 TyrValLeuGlyAlaTyrAlaLysThrLeuSerAspValPheSerAlaAlaPheCys 593  
 Qy 2734 GAAGTATTGTAGATCTAAAGATTATGTTAGTGTCTGCTTCCATCATCATGCTTGCGATA 2793  
 Db 594 GlnLeuPheGlyArgAspLysAspTyrLeuValSerLysAsnAsnSerAsnIleTyrAla 613  
 Qy 2794 GATCCGTTTATCTATCTACCAACAAGCTTTATGTCGATCTTATGTCGAGATCG 2853  
 Db 614 GlySerIleTyrTyrGlnHisThrSer-----PheTrrpAspAla 626  
 Qy 2854 TTT-----ATCCGT 2862  
 Db 627 TrrpAsnLeuLeuGlnSerThrLeuGlyAlaGlnAlaProLeuValLeuAsnAlaGln 646  
 Qy 2863 CTAAGTACGGTGGTGGGATCAGCATATGAACCTCATATACA-----TTTGCAGAG 2916  
 Db 647 LeuThrTyrSerHisThrSerAsnAspMetLysThrAsnMetThrLysTyrAlaPro 666  
 Qy 2917 GAGACGATGTT-----COTTGGGATATAAATGCTGCTGCTGGAGAG 2958  
 Db 667 GlnGlyValValTyrProGluLysGlyAspTrrpGlyAsnAspCysPheGlyValGlu 686  
 Qy 2959 ATTGAGCGGGATACCATGTTGTTGATTAATCCATTAAGCTCTATTTGAATGAGTTGGGT 3018  
 Db 687 LeuGlyAlaThrValProIleGluSerProTyrSerSerLeuPhe---AspMetTyrSer 705  
 Qy 3019 CCTTTGCTGCAAGCTGAGTTTCTTATCCGATCATGATCTTTTACAGAGGAGCGCAT 3078  
 Db 706 PropheLeuArgPheGlnLeuValTyrAlaHisGlnGluAspPheLysGluAsnAsnSer 725  
 Qy 3079 ---CAAGCTCGGCAATCAAGAGCGGACATCTCTAAATCTATCATGTTCTCTGTTGGAGTG 3135  
 Db 726 ThrGluGlyArgTyrPheGluSerSerAspLeuThrAsnLeuSerMetProIleGlyVal 745  
 Qy 3136 AAGTTTGATCGATGTTCTAGTACATCCTTAATAATATAGCTTTATGCGCGCTTATATC 3195  
 Db 746 LysPheGluArgPheSerAspAsnAspIleAlaSerTyrAsnValThrLeuAlaTyrAla 765  
 Qy 3196 TGTGATGTTATCGACCATCTCTGGTACTGAGACAAGCTCTCTATCCATCAAGAGACA 3255  
 Db 766 ProAspLeuValArgSerAsnProAspCysLysThrSerLeuLeuValSerProThrThr 785  
 Qy 3256 -----TGGACAACAGATGCTCTTTCATTTAGCAAGCATGCGATTGTTGGTTAGAGATCT 3309  
 Db 786 AlaValTrrpLeuThrLysAlaThrAsnLeuAlaArgHisAlaPheIleValLysAlaGly 805  
 Qy 3310 ATGTATGCTTCTTAAAGTAATATAGATATATAGATATATAGGCAATGGAAGATATGATATCGA 3369  
 Db 806 AsnTyrLeuSerLeuSerSerAsnPheGluIlePheSerGlnPheGlyPheGluLeuArg 825  
 Qy 3370 GATGCTTCTCGAGGCTATGGTTGATGTCAGGAGTAGAGTCGGGTC 3417  
 Db 826 GlySerSerArgThrTyrAsnValAspLeuGlySerLysIleGlnPhe 841  
 RESULT 33  
 ID P71132\_CHLAB PRELIMINARY; PRT; 847 AA.

AC P71132; OS:SL6J4;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 13-FEB-2005 (TrEMBLrel. 31, Last annotation update)  
 DE POMP91A (Polymorphic outer membrane protein).  
 GN Name=pmp13G; Synonyms=pomp91A; OrderedLocusNames=CAB281;  
 OS Chlamydomonas abortus.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiales.  
 OX NCBI\_TaxID=83555;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=S26/3;  
 RC MEDLINE=98187897; PubMed=9529048;  
 RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;  
 RT "Molecular cloning and characterization of the genes coding for the  
 RT highly immunogenic cluster of 90-kilodalton envelope proteins from the  
 RT Chlamydia psittaci subtype that causes abortion in sheep.";  
 RL Infect. Immun. 66:1317-1324(1998).  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=S26/3;  
 RC MEDLINE=96406378; PubMed=8810511; DOI=10.1016/0378-1097(96)00281-9;  
 RA Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;  
 RT "Identification of a multigene family coding for the 90 kDa proteins  
 RT of the ovine abortion subtype of Chlamydia psittaci.";  
 RL FEBS Microbiol. Lett. 142:277-281(1996).  
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RP STRAIN=S26/3;  
 RC PubMed=15837807; DOI=10.1101/gr.3684805;  
 RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,  
 RA Livingstone M., Cerdano-Tarraga A.-M., Harris B., Doggett J.,  
 RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,  
 RA Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.;  
 RT "The Chlamydomonas abortus genome sequence reveals an array of  
 RT variable proteins that contribute to interspecies variation.";  
 RL Genome Res. 15:629-640(2005).  
 DR EMBL; U65942; AAC15921.1; -; Genomic DNA.  
 DR EMBL; CR848038; CAH63731.1; -; Genomic DNA.  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR InterPro; IPR005546; Auto transportbeta.  
 DR InterPro; IPR006315; Auto transporter.  
 DR InterPro; IPR011427; ChlamPMP\_M.  
 DR InterPro; IPR003368; Chlamydia\_PMP.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF02415; ChlamPMP\_M; 1.  
 DR Pfam; PF07548; ChlamPMP\_M; 1.  
 DR TIGRFAMs; TIGR01414; autotrans bar1; 1.  
 DR TIGRFAMs; TIGR01376; POMP\_repeat; 3.  
 KW Complete proteome.  
 SQ SEQUENCE 847 AA; 90695 MW; 754C958E7F11798 CRC64;  
 Alignment Scores:  
 Pred. No.: 2,77e-54 Length: 847  
 Score: 950.00 Matches: 282  
 Percent Similarity: 41.5% Conservative: 145  
 Best Local Similarity: 27.4% Mismatches: 332  
 Query Match: 12.1% Indels: 270  
 DB: 2 Gaps: 33  
 US-10-701-844-1 (1-4435) x P71132\_CHLAB (1-847)  
 Qy 490 TACGATGGGAGAGCGTTAACTGATGATATTTCCCTTATCTGTTATAGGAGATCCGAGTGG 549  
 Db 36 TyrAsnGlyAsnThrAsnSerGluProPheAsnProLeuSerThrSerAsnSerAsnGly 55  
 Qy 550 ACTACTGTTTCTCGAGGAGGAGTTAACATTA-----AAAATCTTGCAATCTTATT 603  
 Db 56 ThrIleTyrThrCysThrGlyAsnIleAlaTyrAlaGlyLeuAspGlySer--- 74  
 Qy 604 GCAGCTTTGCTTAACTGTTTGGGAACTTATAGGAGATTTTACTGTTTATAGGAGAGA 663  
 Db 75 ---GlyLeuSerSerSerCysPheThrAspThrAlaGlyAsnLeuSerPheLeuGlyAsn 93







Db 116 ThrLeuGlyLeuSerGlyPheSerLeuPheSerCysAlaTyrCys----- 130  
 QY 799 GTACTGCTGCTGCAACGACTAATAAGGTAGC---CAGACTCCGACGACAACATCTACA 855  
 Db 131 -----ProProGlyThrThrGlyTyrGlyAlaIleGlnThrLysGlyAsnThrThrLeu 148  
 QY 856 CCGTCTAATGGTACTATTATTCTTAAACAGATCTTTTGTACTCAATAATGAGAAGTTC 915  
 Db 149 LysAspAsnSerSerLeu----- 154  
 QY 916 TCATTCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACG 975  
 Db 154 ----- 154  
 QY 976 GTTCAAGGAATTAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGCA 1035  
 Db 155 -----ValPheHisLeuAsnCysSerThrAlaGluGlyGly 166  
 QY 1036 GCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAACGAGGCTCTATTGCTTTGTA 1095  
 Db 167 AlaIleGln----- 169  
 QY 1096 GCGAATGTTGCAGGAGTAAGAGGGGAGGAGTTGCTGCTTCCAGGATGGCGACAGGGA 1155  
 Db 169 ----- 169  
 QY 1156 GTGTCATCATCTACTTCAACAGAGAATCCAGTAGTAGTTTTCAGGAATACTGCGGTA 1215  
 Db 169 ----- 169  
 QY 1216 GAGTTTATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCT 1275  
 Db 169 ----- 169  
 QY 1276 TTCCTGAATAATGMAAAACCTTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTGCT 1335  
 Db 169 ----- 169  
 QY 1336 GCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATACGAGATGGAGGA 1395  
 Db 169 ----- 169  
 QY 1396 GCTATCTCTGTAGAATGTGGCGACAGCAGGA-----TCCATAACTCTGGATCA 1446  
 Db 170 -----CysLysGlySerSerAspAlaGluLeuLysIleGluAsnAsnGlnAsnLeu 186  
 QY 1447 GTTTCCTTTGATGGAGGGAGTAGTTTCTTTAGTAGCAATGTAGCTGTGGGAAAGGG 1506  
 Db 187 Val-----PheSerGluAsnSerSerThrSerLysGly 197  
 QY 1507 GGAGCTATTATGCAAAAAGCTCTCGTTTGCTAACTGTGCGCCCTGTACAAATTTTAAAG 1566  
 Db 198 GlyAlaIleTyrAlaAspLysLeuThrIleValSerGlyGlyProThrLeuPheSerAsn 217  
 QY 1567 AATATCGCTAATGAT-----GGTGGAGCGATTTATTATTAGAGAA--TCT 1608  
 Db 218 AsnSerValSerAsnGlySerSerProLysGlyAlaIleSerIleLysAspSerSer 237  
 QY 1609 GGAGGCTCAGTTTATCTGCTGATTATGAGATATTATTTTCGATGGGAAT---CTTAAA 1665  
 Db 238 GlyGluCysSerLeuThrAlaAspLeuGlyAspIleThrPheAspGlyAsnLysIleIle 257  
 QY 1666 AGAAGCAGCAAGAGAGATGCTGCGGATGTTAATGCGGTAACTGTGCTCTCAAGCCATT 1725  
 Db 258 LysThrSerGlyLysSerSer-----ThrValThrArgAsnSerIle 271  
 QY 1726 TCGATGGATCGGAGGGAATAACGACATTAGAGATTAAGAGGAGGCGCATCAGATTCTC 1785  
 Db 272 AspLeuGlyThr---GlyLysPheThrLysLeuArgAlaLysAspGlyPheGlyIlePhe 290  
 QY 1786 TTTAATGATCCATCGATGCGCAACCGAATAACGAGCGGCGAGCTCTTCCAAACTT 1845  
 Db 291 PheTyrAspProIle---ThrGlyGlySerAspGlu----- 302

QY 1846 CTAAAAATTAACGATGGTGA-----GGATACACAGGGGATATGTTTTTCT----- 1893  
 Db 303 LeuAsnIleAsnLysLysGluThrValAspTyrThrGlyLysIleValPheSerGlyGlu 322  
 QY 1894 -----AATGAAGCAGTACTTTGTACCAAAAT 1920  
 Db 323 LysLeuSerAspGluLysAlaArgAlaGluAsnLeuAlaSerThrPheAsnGlnPro 342  
 QY 1921 GTTACGATAGACAGGAAGGATTTCTTCTGTAAGAAAGCGCAAAATTTACGTGATTTCT 1980  
 Db 343 IleThrLeuSerAlaGlySerLeuValLeuLysAspGlyValSerValThrAlaLysGln 362  
 QY 1981 CTAACTGACACAGGTGGAGT---CTGTATATGGAAGCTGGAGTACATGGGATTTGTA 2037  
 Db 363 ValThrGlnGluAlaGlySerThrValValMetAspLeuGlyThrThrLeu----- 379  
 QY 2038 ACTCACAAACACCAACAGCCTCTCGCGCTAATCAGTTGATCAGCTGCTTCCATCTCG 2097  
 Db 380 -----GlnThrProSerSerGlyGlyGluThrIleThrLeuThrAsnLeu 394  
 QY 2098 CATTTGCTCTTTCTTTTGTAGCAAAACATGCAATGCAATGCAATCTCTACCAATCCT 2157  
 Db 395 AspIleAsnIleAlaSerLeuGlyGlyGlyThrSer----- 408  
 QY 2158 CCAGCGCAAGATCTCATCTGCAGTCAATTGGTAGCACAACTGCTGGTTCTGTTACAAT 2217  
 Db 409 -----ProAlaLysLeuAlaThrAsnThrAlaSerGlnAlaIleThr 422  
 QY 2218 AGTGGCCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGTAGTATGGCTA 2277  
 Db 423 IleAsnAlaValAsnLeuValAspAlaAspGlyAsnAlaTyrGlu-----AspProIle 440  
 QY 2278 GGTCTTAATCAAAATCAATGCTCTGTAATATACAGTTAGGAGTAAAGCCCGAGTAAT 2337  
 Db 441 LeuAlaThrSerLysProPheThrAlaIleValAlaThrThrAsnAlaSerThrValThr 460  
 QY 2338 GCCCATCAGATTGACTCTAGGGAATGAGTGCCT-----AAGTATGGCTATCAAGGA 2391  
 Db 461 GlnProThrAsp---AsnLeuThrAsnTyrValProThrHisTyrGlyThrGlnGly 479  
 QY 2392 AGCTGAAGCTTCGCTGGGATCCTAATACAGCAAAATAATGGTCTTATATCTCTGAAGCT 2451  
 Db 480 AsnTyrThrValThrTrpAspThrGluThrAlaThrLys-----ThrAlaThrLeu 496  
 QY 2452 ACATGCACTAAACTGGGTATATCTCGGCGCTGAGCAGTAGCTTCTTTGGTTCCAAAT 2511  
 Db 497 ThrTrpGluGlnThrGlyTyrSerProAsnProGluArgGlnGlyProLeuValProAsn 516  
 QY 2512 AGTTTATGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATTCAGCAAGTGTG 2571  
 Db 517 ThrLeuTrpGlyAlaPheSerAspLeuArgAlaIleGlnAsnLeuMetAspIleSerVal 536  
 QY 2572 GATGGGCGCTCTTTATTTGTCAGGATTTATGGTTTCTGGAGTTTCGAATTTCTTCTATCAT 2631  
 Db 537 AsnGlyAlaAspTyrHisArgGlyPheTrpValSerGlyLeuAlaAsnPheLeuHisLys 556  
 QY 2632 GACCGGATGCTTTAGTACGAGGATTCGCTATATAGTGGGGGTATTCCTTAGGAGCA 2691  
 Db 557 SerGlySerAspThrLysArgLysPheArgHisAsnSerAlaGlyTyrAlaLeuGlyVal 576  
 QY 2692 AACTCTACTTTGGATCA---TCGATGTTTGGTCTAGCATTTACCGAAGTATTGTTGTA 2748  
 Db 577 TyrAlaLysThrProSerAspAspIlePheSerAlaAlaPheCysGlnLeuPheGlyLys 596  
 QY 2749 TCTAAAGATATGATGATGCTCGTTCCAAATCATCATCTGCTGATAGGATCCGTTTATCTA 2808  
 Db 597 AspLysAspTyrLeuValSerLysAsnAsnAlaAsnIleTyrAlaGlySerLeuTyrTyr 616  
 QY 2809 -----TCTACCAACAGCTTTTATGTCGATTCCTATTGTTCCGGAGAT 2850  
 Db 617 GlnHisIleSerTyrTrpSerAlaTrpGlnAsnLeuLeuGlnAsnThrIleGlyAlaGlu 636

```

QY 2851 GCGTTATC-----CGTGCTAGCTACGGGTTTGGATCAGCATCATGAAACC 2898
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
637 AlaProLeuValLeuAsnAlaGlnLeuThrTyrcysHisAlaSerAsnAspMetLysThr 656
QY TCATAT-----ACATTTGCAGAG---GAGAGCGCATGTT 2928
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
657 AsnMetThrThrThrTyAlaProArgLysThrThrTyAlaGluLeuLysGlyAsp--- 675
QY CGTTGGGATAAATACGTCTCGCTGGAGAGATGGAGCGGATACCAATGATGATTA 2988
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
676 --TpgLysAsnAspCysPheGlyValGluLeuGlyAlaThrValProIleGln---Thr 693
QY CCATCAGCTCTATTGGAATGAGTTCGTCCTTCGTCGCAAGCTGAGCTTCTTATGCC 3048
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
694 GluSerSerLeuLeuPheAspMetTyrSerProPheLeuLysPheGlnLeuValHisThr 713
QY GATCATGATCTTTTACAGAG---GAAGCGCATCAAGCTCGGGCATTCAGAGCGGCAT 3105
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
714 HisGlnAspAspPheLysGluAsnAsnSerAspGlnGlyArgTyrPheGluSerSerAsn 733
QY CTCCTAAATCATCATAGTTCCTGTGGAGTGAAGTTTGATCGATGTTCTAGTACATCCT 3165
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
734 LeuThrAsnLeuSerLeuProIleGlyIleLysPheGluArgPheAlaAsnAsnAspThr 753
QY AATAAATATAGTTTATCGCGCTTATCTGTGATGCTTATCGCACCATCTCTGGTACT 3225
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
754 AlaSerTyHisValThrAlaAlaTySerProAspIleValArgSerAsnProAspCys 773
QY GAGACCAACGCTCTCATCCATCAAGAGACA-----TGGACAACAGATGCGCTTCATTTA 3279
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
774 ThrThrSerLeuValSerProAspSerAlaValThrPvalThrLysAlaAsnAsnLeu 793
QY CAAGACATGAGTGTGTAGAGATCTATGATGCTTCTTCAACAAGTAATATAGAA 3339
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
794 AlaArgSerAlaPheMetLeuGlnAlaGlyAsnTyLeuSerLeuSerHisAsnIleGlu 813
QY GTATATGCCATGGAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGTGATGCA 3399
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
814 IlePheSerGlnPheGlyPheGluLeuArgGlySerSerArgThrTyAsnValAspLeu 833
QY GAAAGTAGATCCGGTTC 3417
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
834 GlySerLysIleGlnPhe 839

RESULT 35
PMP2 CHLPN STANDARD; PRT; 841 AA.
AC Q923A1; Q9RB73;
DT 16-OCT-2001 (Rel. 40, Created)
DE Probable outer membrane protein pmp2 precursor (Polymorphic membrane
DE protein 2) (Outer membrane protein 7).
GN Name=pmp2; Synonyms=omp7;
GN OrderedAccession=CP0013; CP0761; CPB0015/CPB0016;
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity";
RL Am. Heart J. 138:8491-8495(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=CWL029;
RC MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;

```

```

RT *Comparative genomes of Chlamydia pneumoniae and C. trachomatis. ";
RL Nat. Genet. 21:385-389(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -1- CAUTION: Ref.4 and Ref.5 sequences differ from that shown due to a
CC frameshift in position 673.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ133035; CAB37083.1; -; Genomic DNA.
CC EMBL; AE001586; AAD18172.1; -; Genomic DNA.
CC EMBL; AE002235; BAF38561.1; -; Genomic DNA.
CC EMBL; BA000008; BAA98223.1; ALT_FRAME; Genomic DNA.
CC EMBL; AE017157; AAF97948.1; ALT_FRAME; Genomic DNA.
CC EMBL; AE017157; AAF97949.1; ALT_FRAME; Genomic DNA.
CC FIC; E72130; E72130.
CC TIGR; CP0761; -.
CC InterPro; IPR005546; Auto_transp_beta.
CC InterPro; IPR011427; ChlamPMP_M.
CC InterPro; IPR003368; Chlamydia_PMP.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02415; Chlam_PMP; 4.
CC Pfam; PF07548; ChlamPMP_M; 1.
CC TIGR; TIGR01376; POMP_repeat; 3.
CC Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
KW SIGNAL
FT CHAIN 25 841 Probable outer membrane protein pmp2.
FT CONFLICT 784 784 N -> S (in Ref. 4).
SQ SEQUENCE 841 AA; 89601 MW; 9064D60D0678D24C CRC64;
Alignment Scores:
Pred. No.: 1.38e-53 Length: 841
Score: 939.50 Matches: 289
Percent Similarity: 39.5% Conservative: 134
Best Local Similarity: 27.0% Mismatches: 358
Query Match: 11.9% Indels: 291
DB: 1 Gaps: 30

```

US-10-701-844-1 (1-4435) X PMP2 CHLPN (1-841)

Qy	382	ATGCCAAACGCTTTTCCATAAAGTTCTTTCTTTCAATGATTCAGCTTATTTCTTGTGCTGCTCT	441
Db			
Qy	1	MetLysileProLeuAargPheLeuLeuLeuLeuLeuValProThrLeuSerMetSerAen	20
Db			
Qy	442	TTAAATCGGGGGGATATGCAGCAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG	501
Db			
Qy	21	LeuLeuGlyAlaAlaThrThrGluGluLeuSerAlaSerAenSer---	39
Db			
Qy	502	ACGTTAACTGTATCATTTCCATATCTACTGTTATAGGAGATCCGAGTGGGACTACTGTTTT	561
Db			
Qy	40	ThrSerThrThrSerPheSerSerLysThrSerSerAlaThrAspGlyThrAenTyrVal	59
Db			
Qy	562	TCTCGAGGAGGTTTAAACATTAAATAAAATCTTTGACAAATCTATATGCGACTTTGCTTTAAAT	621
Db			
Qy	60	PheLysAspSerValValIleGluAenValProLysThrGlyGluThrGlnSerThrSer	79
Db			
Qy	622	TGTTTTGGGAAC-----TTATTAGGAGTTTACTGTTTATAGGAGATCCGAGTGGGAGACACTCGTTG	675
Db			
Qy	80	CysPheLysAenAspAlaAlaAlaGlyAspLeuAenPheLeuGlyGlyPheSerPhe	99
Db			
Qy	676	ACTTTTCGAGAACATA---CGGACTTCTACAAATGGGGCAGCTCTAAATGATTAATCGCTGCT	732
Db			
Qy	100	ThrPheSerAenIleAspAlaThrThrAlaSerGlyAlaAlaIleGlySerGluAlaAla	119
Db			
Qy	733	GATGACTGTTTACTATTGAGGGTTTAAAGAAATTATCTTTTCCAATTGCAATTCATTA	792
Db			
Qy	120	AenLysThrValThrLeuSerGlyPheSerAlaLeuSerPhe-----	133
Db			
Qy	793	CTTGCCGTACTGCTGTCGCAACACTAATAAGGTAGCCAGACTCCGACGACCAATCT	852
Db			
Qy	134	-----LeuLysSerProAlaSerThrVal	141
Db			
Qy	853	ACACGCTCTAATGGTACTATTATTCTAAACACAGACTTTTGTGTACTCAATAATAGAAAG	912
Db			
Qy	142	ThrAenGlyLeuGlyAlaIleAenValLysGlyAenLeuSerLeuLeuAenAspAenLys	161
Db			
Qy	913	TTCTCATTTCTATAGTAATTTAGTCTCTCGGAGATGGGGAGCTATAGAT---	969
Db			
Qy	162	ValLeuIleGlnAenAspAenPheSerThrGlyAspGlyGlyAlaIleAenCysAlaGlySer	181
Db			
Qy	970	TTAACGGTTCAAGGAATTTAGCAAGCTTTGTCTCTCCAGAAATATCTGCTCAAGCTGAT	1029
Db			
Qy	182	LeuLysIle-----	184
Db			
Qy	1030	GGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCTACGAGGCTCCTATTGCC	1089
Db			
Qy	185	-----AlaAenAenLysSerLeuSer	191
Db			
Qy	1090	TTTGTAGCGAATTTGTCAGGAGTAAGAGGGGAGGGAATGTGCTGTTCCAGGATGGGCAG	1149
Db			
Qy	192	PheIleGlyAenSerSerThrAargGlyAlaIleHisThr-----	206
Db			
Qy	1150	CAGGGAGTGTCATCATCTACTTCTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACT	1209
Db			
Qy	206	-----	206
Db			
Qy	1210	CGGGTAGATTGTATGGGNAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC	1269
Db			
Qy	207	-----LysAen	208
Db			
Qy	1270	GTTGCTTTCTCGAATAATGGAAAAACCTGTTTCTCAACAATGTTGCTTCTCTCTGTTTAC	1329
Db			
Qy	209	LeuThrLeuSerSerGlyGlyGluThrLeuPheGlnGlyAenThrAla-----	224
Db			
Qy	1330	ATTGCTGCTAAGCAACCAACAAAGTGGACAGGCTTCTAATACGAGTAATAATTACGGAGAT	1389
Db			
Qy	225	-----ProThr-----	226
Db			
Qy	1390	GGAGGAGCTACTCTTCTGTAAAGAAATGGTGGCAGCAGGATCCAATAACTCTGGATCAGTT	1449
Db			
Qy	226	-----	226
Db			

1450 TCCTTTGATGGAGGAGGTAGTTTTCTTTAGTAGCAATGCTACTCTCGGAAGGGGA 1500  
Db -----AlaLaGlyLys----- 230

QY GCTATTATTGCCAAABAGCTCTCGTGTGTAACTGTGGCCCTGTACAAATTTTAAAGGAAT 1569  
Db ----- 230

QY ATCGCTAATGATGGTGAGCGAATTAATTTAGAGAATCTGGAGAGTCAGTTTATCTGCT 1629  
Db ::::: 1630

QY GlyGlyAlaIleAlaLeuSerGlyThrLeuSerIleSerGly 246  
Db -----GlyGlyAlaIleAlaLeuSerGlyThrLeuSerIleSerGly 246

QY CATTATGGAGATATATTTTCGATGGGATCTTAAAGAACAGCACAAGAGATGCTGCC 1689  
QY AspSerGlyAspIleIlePheGluGly-----AsnThrIle 258  
Db

QY CATGTTAATGCCTAACTGTCTCTCACAGCCATTTTCATGGATCGGAGGGAATA 1749  
Db GlyAlaThrGly---ThrValSerHisSerAlaIleAspLeuGlyThrSerAlaLysIle 277

QY ACGCATTAAGAGCTAAAGCAGGCGCATCAGATCTCTTAAATGATCCATCGAGATGGCA 1809  
Db ThrAlaLeuArgAlaGlnGlyHisThrIleTyrrPheTyrrAspProIleThrVal--- 296

QY AACCGAATATACCAGCCAGCCAGCTCTTCCAACCTCTTAAAAATTAAC- 1857  
Db ThrGlySerThrSerValAlaAspAla-----LeuAsnIleAsnSerProAspThr 313

QY GATGTGAAGNATACACAGGGGATATGTTTTGCT- 1893  
Db GlyAspAsnLysGluTyrrThrGlyThrIleValPheSerGlyGluLysLeuThrGluAla 333

QY AATGAGCAGCATCTTTGTACCAAATATGTTACGATAGACAA 1935  
Db GluAlaLysAspGluLysAsnArgThrSerLysLeuLeuGlnAsnValAlaPheLysAsn 353

QY GGAAGGATGTTCTTCGTGAAGGCAAAATATCATGAGTAATCTCTTAAGTCAGACAGGT 1995  
Db GlyThrValValLeuLysGlyAspValValLeuSerAlaAsnGlyPheSerGlnAspAla 373

QY CTGTATATGAAGCTGGGAGTACATGGGATTTGTAACTCCACCAACCACCA 2052  
Db AsnSerLysLeuIleMetAspLeuGlyThrSerLeu----- 385

QY CAACAGCTCTCGCGCTAATCATGTTGATCAGCTTCCCAATCTGCAATTTGCTCTTCT 2112  
Db ValAlaAsnThrGluSerIleGluLeuThrAsnLeuGluIleAsnIleAsp 402

QY TCTTTGTAGCAAAACAATGCTAGTTACGAATCTCTTACCAATCTCCAGCGCAAGATTCT 2172  
Db SerLeuArgAsnGlyIleValylIle----- 410

QY CATCTGCGAGTCATTTGGTAGCACAACTGCTGTTCTGTACAAATTAGTGGCGCTATCTTT 2232  
Db LysLeuSerAlaAlaThrAlaGlnLysAspIleArgIleAspArgProValVal 428

QY TTGAGGATTCGATGATACAGCTTATGATAGG- 2266  
Db LeuAlaIleSerAspGluSerPheTyrrGlnAsnGlyPheLeuAsnGluAspHisSerTyrr 448

QY GATTGCTAGGTTCTTAATCAAAAATCAATGCTCGAAATTAACAGTTAGGAGCTAAGCCC 2322  
Db Asp-----GlyIleLeuGluLeuAspAlaGlyLysAspIle 460

QY CCAGCTTAATGCCCCATCAGATTTGATCTTAGGGAATGAGATCGCTTAAGTATGCGCTATCAA 2388  
Db ValIleSerAlaAspSerArgSerIleAspAlaValGlnSerPro---TyrrGlyTyrrGln 479

QY GGAAGCTGGAAGCTTCGTTGGGATCCTAATACAGCAAAATATGGTTCCTTACTCTGAAA 2446  
Db GlyValThrIleAsnThrSerThrAspAspLys-----Lys 492



Db	127	SerCysAlaTyrCys-----CysProProGlyThrThrGlyTyrGly	140
Qy	829	AGC---CAGACTCCGACGACACATCTACACCGTCTAAATGGTACTATTATTCTAAACA	885
Db	141	AlaIleGlnThrLysGlyThrThrThrLeuLysAspAsnSerSerLeu	156
Qy	886	GATCTTTTGTACTCAATAATGAGAAGTCTCATTTCTATAGTAAATTTAGTCTCTGGAGAT	945
Db	156	-----ValPhe	156
Qy	946	GGGGAGCTATAGATGCTAAAGAGCTTAA CGGTTCAAGGAATTAGCAAGCTTTTGTGTTCTTC	1005
Db	157	-----ValPhe	158
Qy	1006	CAAGAAATACTGCTCAAGCTGATGAGGGAGCTTGTCAAGTAGTCAACAGTCTCTGCT	1065
Db	159	HisLysAsnCysSerThrAlaGluGlyAlaIleGln	171
Qy	1066	ATGGCTAACGAGGCTCTTATTTGTCCTTTGTAGCGAATGTTTCGAGGATGAGAGGGGGG	1125
Db	171	-----	171
Qy	1126	ATTGCTGCTGTTCAGATGGCCAGGAGTGTCAATCTTCTCAACAGAGAATCCA	1185
Db	172	-----CysLysSerSerSerThrAlaGlu	180
Qy	1186	GTAGTAACTTTTCCAGAAATACTGCGGTAGAGTTTGATGGGAACGTAGCCGAGTAGGA	1245
Db	180	-----	180
Qy	1246	GGAGGATTTACTCTCGGAACTGTTTCTCGAATATGAGAAACCTTGTTCCTC	1305
Db	181	-----LeuLysLeuGluAsnLysAsnLeu	189
Qy	1306	AACAATGTTGCTTCTCTGTTTACATGCTCTAAGCAACCAACAAGTCGACAGGCTTCT	1365
Db	189	-----	189
Qy	1366	AATACGAGTAATAATTACGGAGATGGAGGAGCTATCTTCTAAGAATGGTGGCAAGCA	1425
Db	189	-----	189
Qy	1426	GGATCCAATACTCTGGATCAGTTTCTTGTATGAGAGGAGTAGTTTCTTTAGTAGC	1485
Db	190	-----ValPheSerGlu	193
Qy	1486	ATGTAGTCTGCGGAAAGGGGAGCTATTATGCAAAAAGCTCTCGTGTGCTAACTGT	1545
Db	194	AsnSerSerLysGluLysGlyAlaIleTyrAlaAspLysLeuThrAlaAsnLeuGlyAspIleThr	213
Qy	1546	GSCCTGTACAAATTTTAAAGGAATATCGCTAAT-----GATGGTGGAGCG	1590
Db	214	GlyProThrLeuPheSerAsnAsnSerValSerHisAsnSerSerProLysGlyGlyAla	233
Qy	1591	ATTATTATTAGGAGAATCT---GGAGAGCTCAGTTTATCTGCTGATATGAGAGATATTATT	1647
Db	234	IleCysIleLysAspSerAspGlyGluCysSerLeuThrAlaAsnLeuGlyAspIleThr	253
Qy	1648	TTGATGGGAATCTTAAAGAAACAGCCAAAGAGAATGCTCGGATGTTTAATGGCGTAAC	1707
Db	254	PheAspGlyAsn-----LysIleIleThrThrAsnGlyLysSerProThr	268
Qy	1708	GTGTCCTCACAGCCATTTTCGATGGATGGGAGGGAATAAATACGACATTAAGAGCTAAA	1767
Db	269	ValThrArgAsnSerIleAspLeuGlySerGlyLysPheThrLysLeuAsnAlaLys	288
Qy	1768	GCAGGGCATCAGATCTCTTTAATGATCCATCGAGATGGCAACGAAAGGAAATACACGCCA	1827
Db	289	GluGlyPheGlyIlePhePheTyrAspProIleAlaAsnThrGlySerThrGlu	307
Qy	1828	GGCAGCTCTCCAAACTCTTAAATAATTAACGATGTGAGGATACACAGGGGATATTGTT	1887
Db	308	-----IleGluLeuAsnLysThrGluSerAspThrThrThrThrGlyLysIleVal	324

QY	1988	TTT-----GCTAATCGACAGTACT	1908
Dy	325	PheSerGlyGluLysLeuSerAspGluGluYserThrValProAlaAsnLeuLysSeryr	344
QY	1909	TTGTACCAAAATGTTACGATAGACGAAGAAGATTGGTTCTTCGTGAAAGGCCAACAAATA	1968
Dy	345	PhelysGlnProLeuLysIleGlyAlaGlySerLeuValLeuLysAspGlyValThrLeu	364
QY	1969	TCAGTGAAATCTCTAAGTCAGACAGGTGGAGT---CTGTATATGCAAGCTGGAGTACA	2025
Dy	365	GluAlaLysLysIleThrGlnThrLysGlySerThrValValMetAspLeuGlyThrThr	384
QY	2026	TGGGATTTTGTAACTCCACCAACACCACCAACGACCTCCTGCCTCATTAATCAGTTGATCACG	2085
Dy	385	Leu-----GlnThrProSerSerSerGlyGluThrIleThr	396
QY	2086	CTTTCCAATCTGCATTTGTCCTCTTCCTTCCTTTGTTTAGCAACAATCGAGTTAGCATCTCT	2145
Dy	397	LeuthrAsnLeuAspIleAsnIleAlaSerLeuGlyGlyGlyGlyThrAla----	414
QY	2146	CCTACCAATCTCCAGCGCAAGATTCTCATCTCGACGATCAATGGTAGCAACAATCTGGT	2205
Dy	415	-----PrOAlaLysLeualathraenThrAlaSer	424
QY	2206	TCTGTTACAATTAGTGGGCCTATCTTTTTTGAGGATTTGGATGATACAGCTTATGATAGG	2265
Dy	425	GlnAlaIleSerIleAlaAlaValAsnLeuValAsnThrAspSerAsnThryrGlu---	443
QY	2266	TATGATTGGGTAGTCTTAATCAAANAATCAATGTCCTCGAATATACAGTTAGGACTTAAG	2325
Dy	444	--AspProIleLeuSerAlaSerLys--SerPheSerAlaIleThrAlaThrThrSer	461
QY	2326	CCCCCAGCTAATGCCCATCAGATTTGCACCTCTAGGGAATGAGATGCCT-----AAGTAT	2379
Dy	462	SerSerThrValThrProProGluThrAsnLeuLysasnThyrThrProThrHisThr	481
QY	2380	GGCTATCAAGGAAGCTGGAAGCTTGGCTGGAGTCTTAATACAGCAATAANTGGTCTTAT	2439
Dy	482	GlyTyrgingLyAsnTrpThrValThrTripsylnglySerAlaGln---GluLys	500
QY	2440	ACTCTGAAGTACATGACATGAACCTGAACCTGGGTATAATCCTGGCTGACGACGAGTCTCT	2499
Dy	501	ThrAlaThrLeuThrTrpGlunthrInGlytyrSerProAsnProGluahgvalglySer	520
QY	2500	TTGGTTCCAAATAGTTTATGGGGATCCATTTTAGATATACGATCTGCCTCATCAGCAATT	2559
Dy	521	LeuValProAsnThrLeuTrpGlyAlaPheSerAspThrArgAlaIleGlnAsnLeuMet	540
QY	2560	CAAGCAAGTGTGGATGGGGCTCTTATGTCGAGGATATGGTTCCTGGAGTTTCCGAAT	2619
Dy	541	AepIleSerValAsnGlyAlaAspTy-SerArgGlyPheTrpValSerSerLeuAlaAsn	560
QY	2620	TTCCTTCATATGACCGCGATGCTTTAGGTCAGGATATCGTATATATAGTGGGGTTAT	2679
Dy	561	PhelAuLysSerGlySerAspThrLysarglyspheArghHISerAlaGlyTyR	580
QY	2680	TCCTTAGGAGCAAACTCCTACTTTGGATCA---TCGATGTTTGGCTAGCATTTACCGAA	2736
Dy	581	AlaLeuGlyValTyralaGlnThrProSerAspAspValCySerAlaAlaPheCysGln	600
QY	2737	GTATTGGTAGATCTAAGATNTAGTGTGTCCTCAATCATCTGCTGCATAGGA	2796
Dy	601	LeuPheGlyLysAspLysAspTy-PheValSerLysAsnSerSerThrIleTyralagly	620
QY	2797	TCGGTTTACTA-----TCTACCCCAACAGCTTTTATGTGAGTCTTAT	2838
Dy	621	SerIleTyrrGlnHisIleSerTyrrPaenThrTrpaenThrLeuLeuGlnAsnThr	640
QY	2839	TTGTTCCGGAGATCGCTTATC-----CGTGTACTACTAGGGTTTGGGAATCAG	2886
Dy	641	LeuGIalaglalaProLeuValLeuAsnAlaGlnLeuThrTyrcyehisAlaSerAsn	660

```

QY 2887 CATATGAAACCTCATATACA-----TTTGCAGAGGAGCGATGTT 2928
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 AsnMetLysThrAsnMetThrAsnThrTyrThrProLysAsnValThrProSerGluIle 680
QY 2929 CGT-----TGGGATAAATACTGCTGCTGCGTGGAGAGATTGGAGCGGATTACCGATTGTG 2982
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
681 LysGlyAspTTPGlyAsnAspCysPheGlyValGluPheGlyAlaLysAlaPro----- 698
QY 2983 ATTACTCATCTAAGCTCTATTGTAATGAGTTGGCTCTTTCGTCGAAGCTGAGTTTCT 3042
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
699 IleGluThrAlaSerLeuPheAspMetTyrSerProPheValLysLeuGlnLeuVal 718
QY 3043 TATCGCATCATGAATCTTTTACAGAG---GAAGCGCATCAAGCTCGGCGCATTCAGAGC 3099
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
719 HisAlaHisGlnAspAspPheLysGluAsnAsnSerAspGlnGlyArgTyrPheGluSer 738
QY 3100 GGACATCTCCATAATCTATAGTTCTCTCTGGAGTGAAGTTTGATCGATGTTCTAGTACA 3159
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
739 AsnAsnLeuThrAsnLeuSerMetProIleGlyValLysLeuGluLysPheSerHisLys 758
QY 3160 CATCTTAATAATATAGCTTTATGCGGCTTATATCTGTGATGCTTATCGCACCATCTCT 3219
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
759 AspThrAlaSerTyrAsnLeuThrLeuAlaTyrAlaProAspIleValArgSerAsnPro 778
QY 3220 GGTACTGAGACACGCTCTTA-----TCCATCAAGAGACATGCACACAGATGCCCTTT 3273
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
779 AspCysThrAlaSerLeuLeuValSerProThrSerAlaValTrpValThrLysAlaAsn 798
QY 3274 CATTTAGCAAGACATGGAGTTGTGTAGAGGATCATGTATGCTTCTTAACAAGTAAT 3333
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
799 AsnLeuAlaArgHisAlaPheIleLeuGlnAlaGlyAsnTyrLeuAlaLeuThrArgAsn 818
QY 3334 ATAGAAGTATATGGCCATATGAGATATAGATATCGAGATGCTTCTCGAGGCTATGTTTG 3393
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
819 ThrGluLeuPheSerGlnPheGlyPheGluLeuArgGlySerCysArgThrTyrAsnIle 838
QY 3394 AGTCAGGAAGTAGATCCGGTTC 3417
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
839 AspLeuGlySerLysIleGlnPhe 846

RESULT 37
Q823X2_CHLVC PRELIMINARY; PRT; 843 AA.
AC Q823X2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family protein/autotransporter,
DE putative.
GN OrderedLocusNames=CCNA00281;
OS Chlamydomydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapple E.K., Khouri H.M., Federova N.B.,
RA Carty H.A., Umavam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Haia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydomydia caviae (Chlamydia peitacci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL, AE016995; AAP05032.1; -, Genomic_DNA.
DR TIGR, CCA00281; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto transpbeta.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.

```

```

DR Pfam; PF07548; ChlamPMP M; 1.
DR Pfam; PF02415; Chlam_PMP; 4.
DR TIGRFams; TIGR01376; POMP_repeat; 3.
KW Complete proteome.
SQ SEQUENCE 843 AA; 91510 MW; 05CCE3764CF3A43 CRC64;

Alignment Scores:
Pred. No.: 1.61e-44 Length: 843
Score: 803.00 Matches: 270
Percent Similarity: 38.1% Conservative: 137
Best Local Similarity: 25.3% Mismatches: 361
Query Match: 10.2% Indels: 300
DB: 2 Gaps: 36

US-10-701-844-1 (1-4435) x Q823X2_CHLVC (1-843)
QY 418 ATTCTAGCTTATCTTCTGCTCTTTAAATGGGGGGGATATGCGAGCAATCATGOTT 477
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ValLeuLeuPheSerSerPheAlaLeuSer-----IleAlaThrGluLeuAla 24
QY 478 CCTCAAGAAATTTACGATGGGAGACGTTAACTGTA----- 513
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25 -----AspAlaAspThrValAsnLeuAlaAlaGlyPheAsnGlySerSer 39
QY 514 -----TCATTTCCCTATACGTGTTATAGGAGATCCGAGTGGGACTGTTTTCTGCA 567
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40 SerGluThrPheAsnValLysGlnThrAsnAsnValGluGlyThrTyrThrLeuThr 59
QY 568 GGAGAGTTAACTATAAAAAATCTTGACAAATCTTATTCAGCTTTGCGCTTTAAAGTTGTTTT 627
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 GlyAlaValSerPheAlaAsnIle---AsnLysPheAspGlnAlaAspThrSerCysPhe 78
QY 628 GGGAACTTATPAGGAGTTTATCTGTTTGGGAGAGGACACTCGTTGACTTTTCGAGAAC 687
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 AlaAsnSerAlaGlyAspLeuThrPheThrGlySerArgLeuLeuTyrPheAsnAsn 98
QY 688 ATACGGACTTCTACAAATGGGCGAGCTTAAGTAATAGCGCTGCTGATGAGCTGTTACT 747
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
99 IleThrSerSerAlaLysGlyAlaIleSerThrThrAla----- 112
QY 748 ATTGAGGGTTTAAAGAAATATCTCTTTTCCAAATTCGAATTCATTACTTTCGCTACTGCT 807
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 -----AspAlaLysThrLeuThrLeuSerGlyCysLeuSerLeuIlePheTyrMet--- 129
QY 808 GTCGCAACGACTAATAAGGTAGCCAGACTCCGACGACACAATCTACACGCTCTAATGCT 867
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 -----SerProLysGluAspIleGlyAsnGly 138
QY 868 ACTATTATTCTAAAACAGATCTTTTCTTACTCAATAATGAGAAGTTCTCATTTCTATAGT 927
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 AlaIleTyrSerAsnSerMetLeuIleGlnAsnSerAspValSerPheGlyTyr 158
QY 928 AATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAAAGCTTCAAGGAATT 987
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
159 AsnLysSerAlaGlyLysGlySerValIle----- 168
QY 988 AGCAAGCTTTGTGTTCTTCCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTTCAGATA 1047
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 -----LeuCysGluLys 172
QY 1048 GTCAACGATTCTCTGCTATGGCTTAACGAGGCTCTATTCCTTTGTAGGAATGTTGCA 1107
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 SerThrSer----- 175
QY 1108 GGAGTAAGGGGGGAGGATGCTGCTGTTTCAGGATGGCAGGAGGAGTGTCTCATCTCT 1167
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 -----AlaGlyAlaThrSerPro 181
QY 1168 ACTTCAACAGAAATCCAGTAGTAGTAAAGTTTTTCCAGAAATACTGCGGTAGAGTTGATGG 1227
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 ThrLeuThr----- 184
QY 1228 AACGTAGCCCGAGTAGGAGGAGGATTACTCTCTACGGGAACGTTGCTTTCTCTGAATAAT 1287

```



Db 185 -----|||||  
1288 GGAAGAACCTGTTTCTCAACAATGTTGCTCTCTGTTTACATGCTGCTAAGCAACA 1347  
|||  
Db 189 GlyGluPheLeuValLeu-----194  
1348 ACAAGTGGACAGGCTTCTAATACGAGTAATAATTACGGAGTGGAGGAGTATCTTCTGT 1407  
194 -----194  
1408 AAGAATGGTGGCAAGCAGGATCCAACTCTGATCAGTTCCTTTGATGGAGGGA 1467  
194 -----194  
1468 GTAGTGTTCCTTTAGTACCAATGCTGCTGGAAAGGGGAGCTATTATTATCCAAAAAG 1527  
|||  
Db 195 -----AlaAsnLeuSerAlaSerGlyGlyAlaIleTyrAlaGluLys 209  
1528 CTCTCGGTGCTAACTGTGGCCCTGTACNAATTTTAAAGGAATATCGCTAATGAT---GGT 1584  
|||  
Db 210 MetLeuLeuSerAlaGlyGlyAsnThrValPheGlnAlaAsnValThrGlnGluLysGly 229  
1585 GGAGCGATTTATTAGGAATCTGGAGAGCTCAGTTTATCTGCTGATTAATCGAGATATT 1644  
|||  
Db 230 GlyAlaIleAlaIleAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 249  
1645 ATTTTCGATGGGAATCTTAAAGAACACAGCAAGAGAAATGCTGCCGATGTTAATGCCGTA 1704  
|||  
Db 250 ThrPheGluArgAsnIle-----IleIleGluAsnAsnGluThrIleArg-----264  
1705 ACTGTGCTCTCAGACGCCATTTGCGATGGGATCGGAGGGAATAAATACGACATTAAGACT 1764  
|||  
Db 265 -----AsnAlaIleHisLeuGluAspGlyAlaIlePheLeuGlnLeuArgAla 280  
1765 AAGCAGGCGCATCAGATCTCTTAAATGATCCATCGAGATGCAACCGAAATAACACAG 1824  
|||  
Db 281 AlaLysAspArgSerIleTyrPheHisAspProIle---ThrThrGlyAsn-----297  
1825 CCAGCGCAGCTCTCCAACTCTTAAATAATTAAAGATGGTGAAGGATACACAGGGGATATT 1884  
|||  
Db 298 ValAlaAspArgLeuThrLeuAsnAlaAlaAsnGlyAlaThrProTyrGluGlyThrIle 317  
1885 GTTTTGTCTAATGGA-----AGCAGTACTTTGTAC-----1914  
318 ValPheAlaSerGlyValSerTyrValAsnSerProLeuSerLysLeuTyrSerPheSer 337  
1915 CAAATGTTACGATAGCAAGCAAGGATGTTCTCGTGAAGCAAGCAAAATTTATCAGTG 1974  
|||  
Db 338 GlnAspLeuThrLeuAlaAlaGlySerLeuIleLeuLysAspValLeuIleLysAla 357  
1975 AATTCTCTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGAT 2031  
|||  
Db 358 LysSerPheGluGlnAsnGlnLysSerLeuLeuPheMetHisProGlyThr-----374  
2032 TTTGTAACTCCACACACACAGCCCTCTGCGCCTTAATCAGTTGATCAGCTTTCC 2091  
|||  
Db 375 -----ArgLeuGlnThrThrAsnAsnLeu---SerIleLys 385  
2092 AATCTGCAATTTGCTCTTTCTTTGTTAGCAACAATGCACTACGATCCATCCCTCTACC 2151  
|||  
Db 386 AsnLeuHisLeuAspLeuSerLysIleAlaAla---ThrAlaVal-----399  
2152 AATCTCCAGCGCAAGATTTCTATCTCTGCACTCATTTGGTAGCAACAATGCTGTTCTGTT 2211  
|||  
Db 400 -----GluIleAlaAlaThrAlaAspAlaAlaIle 410  
2212 ACAATTAGTGGCCTATCTTTTGTAGGATTTGGATGATACAGCTTATGATAGTATGAT 2271  
|||  
Db 411 GluIleCysGlyProMetValMetHis---ValAspAspGluIlePheTyrAsnGlnGlu 429  
2272 TGCGTAGTCTTAATCAAAAAATCAATGTCCTGAAATTACAGTTAGGACTAAGCCCCCA 2331  
|||

430 AlaLeuAlaAspSerLeuSerPheGluCysLeuHisValArg-----443  
2332 GCTAATGCCCA-----TCAGATTTCAGTCTAGCGAATGAGATGCTCT-----2373  
444 -----AlaProHisLeuAspAsnIleThrVal---AspAspValProLeuIleProIle 460  
2374 -----AAGTATGGCTATCAAGAAAGCTGGAAGCTTGGCTGGATCCT---2415  
461 ThrThrMetGluThrHisArgGlyTyrGlnGlyLysTyrThrValSerTyrGluGluGlu 480  
2416 -----AATACAGCAAAATAATGGTCTCTTATCTCTCAAGCTACATGG 2457  
481 HisGluMetIlePheGlyAsnValSerThrGlnProAsnLysLysMetSerLeuValTyr 500  
2458 ACTAAACCTGGGTATAATCTCT-----GGGCTGAGCGAGTAGCTCTTTGTT 2505  
501 AsnProSerGlyTyrIleProPheValGlyGlyThrGlyGluPheThrThrSerLeuVal 520  
2506 CCAATAGTATTTATGGGATCCATTTTAGATATACGATTCGCATTCAGCAATTCAGCA 2565  
521 ProAsnSerLeuTyrAsnLeuPheLeuAspThrArgPheAlaGlnAlaIleGlu---539  
2566 AGTGTGGATGGCGCTCTTATTTGTCGAGATATGCGTTCTCGAGTTTCGAGTTCTCTC 2625  
540 ---ThrAsnAlaGlnSerProGlyAsnGlyIleTyrIleSerSerLeuThrAsnSerPhe 558  
2626 TATCATGACCGCATGCTTTAGTTCAGGATATCGGTATATTTAGTGGGGTTTATCTCTTA 2685  
559 ArgLysGlySerThrGluAsnAsnHisGlyPheArgHisLysSerSerGlyTyrValAla 578  
2686 GGAGCAAAAC---TCCTACTTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTATT 2742  
579 GlyGlyLysPheGlnThrLeuGlnAspIlePheSerValGlyLysCysGlnLeuPhe 598  
2743 GGTAGATCTAAAGATATTAGTGTGCTTCCAAATCATCATGCTTGCATAGATCCGTT 2802  
599 GlyArgSerLysAspPheGlySerAlaLysSerLysAspLysAlaPheSerGlySerLeu 618  
2803 TATCTATCTACCAACAAGCTTTA-----TGTGGATCTTATTTGTTCCGA-----2847  
619 TyrAlaHisHisSerArgTyrLeuLeuProIleThrArgPheLeuAlaGlyThrSerArg 638  
2848 -----GATGCGTTT 2856  
639 SerGlnProArgPheLeuSerArgIleProLysAspPheProIleAsnPheAspAlaLeu 658  
2857 ATCCGTGCTAGTACGGGTTTCGGAATCAGCATATGAAACCTCATATATCATTTGCAGAG 2916  
659 Ile-----SerTyrSerTyrGlyArgAsnHisMetLysValLysTyrAlaAspHisSer 676  
2917 GAGAGGATGTTCTGGATTAATTAATCTCTGCTGGAGAGATGAGCGGGATTTACCG 2976  
677 GlnThrThrSerSerTyrAsnThrTyrGlyTyrSerAlaGlnIleGlySerLeuPro 696  
2977 ATTTGTACTTACTCCATTAAGCTCTATTGTAATGATTCGCTCTTCTTCCGCAAGCTGAG 3036  
697 CysAlaLeuAspValSerHisThrPhePheGlnTyrValSerProPheValLysLeuHis 716  
3037 TTTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGATTTCAAG 3096  
717 TrpIleTyrAlaHisGlnValGlnPheGlnGlnGlyIleLysArgArgSerPheAsn 736  
3097 AGCGGACATCTCTTAATCTATCAGTTCTCTGTTGGAGTCAAGTTTGCATCGATTTCTAGT 3156  
737 AsnSerAsnLeuLysAsnLeuSerLeuProIleGlyLeuLysIle---GlnGlyGlnSer 755  
3157 ACACATCTCTAATAATATAGCTTTATGGCGCTTATATCTGTGCTTATCTGCTCATCCACCATC 3216  
756 LeuHisHisLeuSerTyrGluLeuThrGlyMetTyrIleAlaAspLeuTyrArgCysAsn 775  
3217 TCTGTGTACTGAGACAACGCTCTCTATCC---CATCAAGAGACATGGAACAACAGATGCTTT 3273  
776 ProGluSerValThrSerLeuIleSerGlyLeuLeuProTyrThrThrAlaAla 795

```
QY 3274 CATTTAGCAACATGAGGTTGTGGTTAGAGGATCTATGTATGCTTCTCAACAGTAAT 3333
D 9 CysCys
D 796 AsnLeuGlyGlnAlaAlaLeuLeuGlnGlySerGlyAsnLeuSerLeuThrSerHis 815
QY 3334 ATAGAGTATATGGCCATGAGGATATGAGTATCGAGATGCTTCTCGAGGCTATGGTTG 3393
D 816 IleAsnIlePheAlaGlnGlyThrValGluPheArgSerSerTyrSerTyrAlaMet 835
QY 3394 AGTCAGGAGTATGAGTCCGTTTC 3417
D 836 AsnAlaGlySerArgValHisPhe 843
```

## RESULT 38

```
PMPI_CHLMU ID PMPI_CHLMU STANDARD; PRT; 867 AA.
AC Q9PL41;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable outer membrane protein pmpl precursor (Polymorphic membrane
protein 1).
GN Name=pmpl; OrderedLocusNames=TC0267;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
```

## NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

```
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Uterback T.R., Berry K.J.,
RA Basse S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwin M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae A39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
(Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC

DR EMBL; AB002294; AAF39136.1; -; Genomic\_DNA.

DR PIR; F81721; F81721.

DR TIGR; TC0267; -.

DR InterPro; IPR005546; Auto transp beta.

DR InterPro; IPR011427; ChlamPMP M.

DR InterPro; IPR003368; Chlamydia PMP.

DR Pfam; PF03797; Autotransporter; 1.

DR Pfam; PF02415; ChlamPMP; 6.

DR Pfam; PF07548; ChlamPMP M; 1.

DR TIGRPFAMs; TIGR01376; POMP repeat; 6.

DR Complete proteome; Membrane; Multigene family; Outer membrane; Signal.

FT SIGNAL 1 29 Potential.

FT CHAIN 30 867 Probable outer membrane protein pmpl.

SQ SEQUENCE 867 AA; 95018 MW; 557994185A9E562 CRC64;

Alignment Scores:

Pred. No.: 1.38e-43 Length: 867

Score: 789.00 Matches: 268

Percent Similarity: 40.5% Conservative: 148

Best Local Similarity: 26.1% Mismatches: 411

Query Match: 10.0% Indels: 200

DB: 1 Gaps: 31

US-10-701-844-1 (1-4435) x PMPI\_CHLMU (1-867)

```
QY 433 TGCTGCTCTTTAAATGGGGGGATATGACGACGAAATCATGGTTCTCAAGGAATTAC 492
D 9 CysCys
QY 493 GATGGGAGAGCTTTAACTGTATCATTTCCCTATCTCTATAGGATCCGAGTGGGACT 552
D 23 PheGlyGlnAsp
QY 553 ACTGTTTTTCTGAGGAGAGTAAACATTAATAAATCTTGACAAATCTTATGACGCTTTG 612
D 42 ValCysThrPheLeuGluAspCysThrMetGluAsnPheSerProAlaLeuLeuSerHis 61
QY 613 -----CCTTTAAGTTGTTTGGGAACCTATTAGGGAGCTTTTACTGTTTA 657
D 62 AlaArgGlnAspProLeuTyrIleIleGlyAsnThr
QY 658 GGGAGAGGACATCGTTGACTTTCGAGACATACGAGACTTCTACAAATGGGGAGCTCTA 717
D 75 -----HisAsnTrpPheValSerAsnLeuHisProSerThrAsnGluGluArgPhe 91
QY 718 AGTAATAGCCGCTGCTGATGACGCTGTTTACTATTAGAGGGTTTTAAAGAAATTCCTTTCC 777
D 92 LeuLysGluLysGlyAsp
QY 778 AATTGCAATTCATTACTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
D 110 AspCysSerSer
QY 838 CCGACGACACATCTACACCGTCTTAATGTTACTATTATTCTTAAACA-----GATCTTTTG 894
D 114 ---SerThrGluAspSerProSer
QY 895 TTACTCAATTAATGAGAGTCTCATCTTCTATGTAATTTAGTCTCTGGAGATGGGGAGCT 954
D 131 LeuArgAsnGlnAsnMetSerPheTyrArgAsnHisSerGluGlySerGlyAla 150
QY 955 ATAGATGCTAAGAGCTTAAACGTTCAAGGAATTAGCAAGCTTTGTGCTTCCCAAGAAAT 1014
D 151 LeuSerThrAspAlaLeuPheLeuGlnHisAsnTyrLeuPheThrAsnPheGluGluAsn 170
QY 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACACAGTCTTCTGCTATGGCTAAC 1074
D 171 SerSerAlaLysGlnGlyAlaIleGlnAlaGlnThr
QY 1075 GAGCTCCTATGCTTTTGTAGCGAATGTTGACAGAGTAAAGAGGGGAGGATGCTGCT 1134
D 183 -----
QY 1135 GTTCAGGATGGGACGAGGAGTGTCTATCTACTTCAACAGAGATCCAGTAGTAAGT 1194
D 184 -----LeuSer
QY 1195 TTTTCCAGAAAT---ACTGGGTAGAGTGTGATGGGAACGTAGCCGAGTAGGAGGAGG 1251
D 186 LeuSerArgAsnValSerSerLeuSerPheSerArgAsnArgAlaAsnLeuAsnGlyGly 205
QY 1252 ATTTACTCTACGAGAACGTTGCTTCTGTAATTAATGGAACCAACCTGTTTCTCAACAAT 1311
D 206 AlaIleCysCysGlnAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 223
QY 1312 GTTGCTTCTCTGTTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1371
D 224 -----Thr
QY 1372 AGTAATAATTACGAGATGGAGGAGCTATCTTCTGTAAGAATGTTGGCAAGCAGGATCC 1431
D 225 AsnAsnSerAlaLeuAsnGlyGlyAlaIleCysCysIleAsnGluGlnAsnLeuSerGlu 244
QY 1432 AATAACTCTGATCAGTTTCTTTGATGAGAGGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491
D 245 LysGlyCysLeuSerLeuAlaTyrAsnGlnGlu-----ThrLeuPheSerGlyAsnSer 262
QY 1492 GCTGCTGGAAAGGGGAGGCTATTATGCAAAAAGCTCTCGGTTGCTTAACGTGGCCCT 1551
```



RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LLG;  
 RA Laroucau K., Souriau A., Rodolakis A.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF243418; AAL36962.1; -; Genomic\_DNA.  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR InterPro; IPR006315; Auto:transporter.  
 DR InterPro; IPR005546; Auto:transporter.  
 DR InterPro; IPR011427; ChlamPMP M.  
 DR Pfam; PF03797; Auto:transporter; 1.  
 DR Pfam; PF07548; ChlamPMP M; 1.  
 DR TIGRFAMs; TIGR01414; autotrans\_bar1; 1.  
 FT NON\_TER  
 SQ SEQUENCE 602 AA; 65561 MW; CA486CFACEC131E2 CRC64;

Alignment Scores:

Pred. No.: 1.57e-43 Length: 602  
 Score: 787.50 Matches: 210  
 Percent Similarity: 48.6% Conservative: 99  
 Best Local Similarity: 33.0% Mismatches: 250  
 Query Match: 10.0% Indels: 77  
 Gaps: 21

US-10-701-844-1 (1-4435) x Q8VU49\_CHLPS (1-602)

QY 1636 GGAGATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTGCCGATGTT 1695  
 DB 2 GlyAspIleThrPheAspGlyAsnLysIleIleThrThrSerArgSerSer----- 19  
 QY 1696 AATGCGTAATGTCCTCCACAGCCATTTCGATGGATCGGAGGGGAAATTAACGACA 1755  
 DB -----ThrValLysArgAsnSerIleSerLeuGlySerGlyGlyPheThrLys 36  
 QY 1756 TTAAGAGCTAAAGCAGGCGCATCTCTTTAATGATCCCATCGAGATGCGAAACGGA 1815  
 DB 37 LeuAsnAlaLysGluGlyPheGlyIlePhePheThrAspProIleAlaAsnThrGlyAsp 56  
 QY 1816 AATAACCGCCAGCGCAGCTCTTCCAAACTCTTAAAAATTAACGATGGTGAAGGA----- 1869  
 DB -----ThrAsnThrGlu-----IleGluLeuAsnLysAlaGluGlyGlySer 70  
 QY 1870 -----TACACAGGGGATATTGTTTTCCT----- 1893  
 DB 71 ThrThrThrThrGlyLysIleValPheSerGlyGlyLysLeuSerAspGluLysLys 90  
 QY 1894 -----AATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACCAAGGAGGATT 1944  
 DB 91 ValAlaAspAsnLeuLysSerThrPheThrGlnProLeuLysIleGlyAlaGlySerLeu 110  
 QY 1945 GTTCTTCTGAAAAGGCCAAATATTACGTGAATTCTTAAGTCACACAGGTGGGAGT--- 2001  
 DB 111 ValLeuLysAspGlyValThrLeuGluAlaLysLysValSerGlnThrAspGlySerThr 130  
 QY 2002 CTGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCACACACCAACAGCCT 2061  
 DB 131 ValValMetAspLeuGlyThr-----LeuGlnThr 141  
 QY 2062 CTGCGCGCTAATCAGTTGATCAGGCTTCCCAATCTGCATTGTCTTCTTCTTTGTTA 2121  
 DB 142 SerSerSerGlyGluThrIleThrLeuThrAsnLeuAspIleAsnValIleSer---Leu 160  
 QY 2122 GCAAAACAATCAGTTTACGAATCTCTACCAATCTCCAGCGGCAAGATTCTCATCTGCA 2181  
 DB 161 GlyGlyGlyGlyValAlaProAspProAlaLysValGluAlaGlnAlaSer----- 177  
 QY 2182 GTCATTGTCAGCAACTCGTGGTCTCTTACAAATTAGTGGCCCTATCTTTTGGAGAT 2241  
 DB 178 -----GlyLysThr-----ValThrIleAsnAla---ValAsnLeuValAsp 190  
 QY 2242 TTGGATGATCAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAAATCAATGTC 2301  
 DB 191 ThrAspGlyAsnAlaLysGlu---TyrProIleLeuAlaThrSerGlnProPheThrAla 209

QY 2302 CTGAAATTACAGTTAGGGACTAAGCCCCAGCTAATGCCCATCAGATTGACTCTTAGGG 2361  
 DB 210 IleIleAlaLysAlaGlySerSerGlyThrThrThrProThrAsp---AsnLeuLys 228  
 QY 2362 AATGAGATGGCT-----AAGTATGGCTATCAAGGAGCTGGAAGCTTGGCTGGATCCT 2415  
 DB 229 AsnTyrThrProThrHisTyrGlyThrGlnGlyAsnThrThrValThrTrpLysLeu 248  
 QY 2416 AATACAGCAAAATATGCTCTGAAATCTTCTGAAAGCTACATGGACTAAACATGGGTATAAT 2475  
 DB 249 GlyThrSerAlaGln---GluGluThrAlaThrLeuThrTrpGluGlnThrAspTyrSer 267  
 QY 2476 CTGGGCTCTGAGGAGTAGCTTCTTGGTCCAAATAGTTATAGGGGATCCATTTTATAGT 2535  
 DB 268 ProAsnProGluArgGlnGlyProLeuValProAsnThrLeuTrpGlySerPheSerAsp 287  
 QY 2536 ATACGATCTCGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTTGCGAGGA 2595  
 DB 288 IleArgAlaIleGlnAsnLeuIleAspIleSerValAsnGlyAlaAspTyrArgArgGly 307  
 QY 2596 TTATGGTCTTCTGAGTTCGAAATTTCTTCTATCATGACCGCATGCTTTAGTGCAGGA 2655  
 DB 308 PheTrpValSerGlyLeuGlyAsnPheLeuHisLysSerGlySerAsnThrLysArgLys 327  
 QY 2656 TATCGTATATTATAGTGGGTTTCTTCTTAGGACAACTCTTACTTTTGGATCA---TCG 2712  
 DB 328 PheArgHisSerAlaGlyTyrAlaLeuGlyValTyrAlaGlnThrSerThrGluAsp 347  
 QY 2713 ATGTTTGGTCTAGCATTCGGAATTTGGTAGATCTAAAGATTATAGTGTGTCGT 2772  
 DB 348 ValPheSerAlaAlaPheCysGlnLeuPheGlyLysAspLysAspTyrPheValSerLys 367  
 QY 2773 TCCAATCATGCTGCTCATAGGATCGCTTATCTA-----TCTACC 2814  
 DB 368 AsnSerSerAsnIleTyrAlaGlySerIleTyrGlnHisIleSerTyrTrpAsnAla 387  
 QY 2815 CAACAGCTTTATGTGATCTCTATTGTTTCGGAGATCGGTTTATC-----CGT 2862  
 DB 388 TrpGlnAsnLeuGlnSerThrIleGlyAlaGlyAlaProLeuValLeuAsnAlaGln 407  
 QY 2863 GCTAGCTACGGTTTGGGAATCAGCATATCAAACTCATATACA----- 2907  
 DB 408 LeuThrTyrCysHisAlaSerAsnAsnMetLysThrAsnMetThrAsnThrTyrValPro 427  
 QY 2908 -----TTTCAGAGGAGAGCGATGTTCTGGTGGATTAATACTCTCTGGCTGGAGAG 2958  
 DB 428 LysAsnValThrLeuSerGluIleLysGlyAspTrpGlyAsnAspCysPheGlyValGlu 447  
 QY 2959 ATTGAGCGGGATTTACCGATTGTGATTTACTCCATCTAAGCTCTATTGTAATGAGTGGCT 3018  
 DB 448 PheGlyAlaMetAlaProIleGlu---ThrProSerSerPheLeuPheAspArgTyrSer 466  
 QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAG---GAAGGC 3075  
 DB 467 ProPheLeuGlnLeuGlnLeuValHisAlaHisGlnAspPheLysGluAsnAsnSer 486  
 QY 3076 GATCAAGCTCGGCATTCAAGACGGACATCTCTTAATATCATCAGTTCTCTTGTGGAGTG 3135  
 DB 487 AspGlnGlyArgTyrPheGluSerSerAsnLeuThrAsnLeuSerMetProIleGlyLe 506  
 QY 3136 AAGTTTGAATCGATTTCTAGTACATCCCTTAATAATATAGCTTTTATGGCGGCTTATATC 3195  
 DB 507 LysPheGluArgPheAlaTyrAsnAspValAlaSerTyrHisLeuThrAlaAlaTyrAla 526  
 QY 3196 TGTGATCTTATCGCACCATCTCTGGTACTGAGACAACGCTCCTA-----TCCCATCAA 3249  
 DB 527 ProAspIleValArgSerAsnProAspCysThrAlaSerLeuLeuValSerProThrSer 546  
 QY 3250 GAGACATGCAACAGATGCTTTTCATTATGACAGACATGGAGTTGTGGTTAGAGATCT 3309  
 DB 547 AlaValTrpValThrLysAlaAsnAsnLeuAlaArgSerAlaPheMetLeuGlnAlaGly 566



```
QY 1804 ATGGCAACGGGAATAACACGACGAGCGAGTCTTCCAACTCTTAAATAATTAACGATGT 1863
Db 291 ValAlaGluGlyThr-----AlaAspSerAen-----LeuGluileAenLysAla 305
QY 1864 GAAGGA-----TACACAGGGGATATGTTTTGCT----- 1893
Db 306 AspGlyGlyThrSerTyThrGlySerilePheSerGlyArgTyThrileGluSerPro 325
QY 1894 -----AATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAACGAAGG 1941
Db 326 HisLysArgMetLysHisValSerThrPheThrGlnProLeuThrLeuSerSerGlySer 345
QY 1942 ATTGTTCTTCGTAAGCGCAAAATATCATGTAATCTCTTAAGTCACAGAGTGGGAGT 2001
Db 346 LeuValLeuGluLysGlyAlaHisLeuLysAlaLysSerLeuThrGlnThrAlaGlySer 365
QY 2002 ---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCACCAACAG 2058
Db 366 LysValileLeuAaspGlnThrSerSerileGlu----- 376
QY 2059 CCTCTCGCGCTAATCAAGTTCATACGCTTTCCAACTGCAATTTGTCCTCTTCTTTG 2118
Db 377 -----ThrLysGluAenileAaspileLysGluLeuTrrPheLysArgLeuAaspPhe 393
QY 2119 TTAGCAAAATGCGAGTTACGATCTCTCTACCAATCTCCAGCGCAAGATTCATCCT 2178
Db 394 -----AenThrProThr-----AlaThr 399
QY 2179 GCAGTCATTTGGTACACAACTGGTCTGTCTACAATTAGTGGGCTTATC---TTTTTT 2235
Db 400 CysIleSerThrSerGlyAenAlaHisThrileThrileLysGlyProLeuGlyValPhe 419
QY 2236 GAGGATTTGGATGATACAGCTTATAGTAGGTAT-----GAT 2271
Db 420 ThrAap---GlnGluThrPheTyAspAenHisAlaLeuAlaTySerileAaspGlnGlu 438
QY 2272 TGGCTAGGTCTTAATCAAAATCAATCTCTGTAATTTACGATTTAGGACGTAAGCCCA 2331
Db 439 PheLeuGlnLeuAlaAaspLys---AaspileThrLysileSerLeuValAaspileProGln 457
QY 2332 GCTAATGCCCATCAGATTGACTCTAGGGAATGAGATCGCTTAAGTAGTGGCTATCAAGGA 2391
Db 458 Ala-----ValArgLysAenLeuAaspSerHisSerGlyTyrglnGly 471
QY 2392 AGCTGGAAAGCTTGGCTGGAT-----CCTAATACAGCAAAATATGGTCTTATCTG 2445
Db 472 LysTrpSerileAaspTrpLysThrValProGlySerThrAenAlaGlyValThrThrLeu 491
QY 2446 -----AAGCTACA-----TGGACTAAACTGGGTATATCTCT-----GGGCT 2484
Db 492 GlyThrLysThrAlaThrValHisTrpArgProThrGlyTyThrileProPheGlyGlySer 511
QY 2485 GAGCGAGTAGCTTCT---TTGGTCCAAATAGTTTATGGGATCCATTTTAGATATACGA 2541
Db 512 GlnGluileThrProLeuValValAenThrLeuTrpGlyAenPheSerAaspileArg 531
QY 2542 TCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGT---CGAGGATTA 2598
Db 532 AenLeuGluArgThrValGlu---SerLeuAlaValAenSerLeuLysSerGluGlyPhe 550
QY 2599 TGGGTTCTCGAGTTTCGAATTTCTTCTATCATGACCGCGATGCTTTAGGTACGGGATAT 2658
Db 551 TrpAlaAlaGlyIleLysAenTyrrLysSerAenSerProAlaGluAenTyrrValPhe 570
QY 2659 CGGTATATTAGTGGGGTTATCTCTAGGACCAACTCTAC---TTTGGATCATCGATG 2715
Db 571 GlnHisHisAenAlaGlyTyrrAlaileGlyMetAenLysHisThrLeuSerGluAenVal 590
QY 2716 TTTGTCATAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTAATAGTGTGTGCTTCC 2775
Db 591 PheSerAlaAlaPheSerGlnLeuPheGlyLysAaspArgPheHisAlaAenGlyHisVal 610
```

```
QY 2776 AATCATCATGCTTGATAGATCGGTTTAT-----CTATCTACCCAA 2817
Db 611 AspHisGlnThrLeuSerGlySerPheTyrrAlaHisValGlySerLeuProMetLeu 630
QY 2818 CAAGCTTTATGTGGATCCTATTGTTGCGAGATGCGTGTATTCGTTGCT----- 2865
Db 631 ArgPheLeuCysGlyGlySerLysAenCysProProGluLeuGlnAlaSerProSerIle 650
QY 2866 -----AGCTACGGGTTTGGGAATCAGCATATGAAACCTCA 2901
Db 651 ProValileValAenAlaGlnLeuSerTyrrSerHisSerAenAenHisLeuThrileHis 670
QY 2902 TATACATTTCCAGAGGAGCGATGTTGTTGGATATAACTCTCTGCTGGCTGGAGATTT 2961
Db 671 HisGluAaspThrThrLysThrThrGlyMetTrpSerAenTyrrSerLeuAlaAlaGluLeu 690
QY 2962 GGACGGGATTTACCGATTTGATTTACTTCACTAAGCTCTATTGTAATGATTCGCTCCT 3021
Db 691 GlySerThrPheValTyrrThrLeuSerLysCysProSerileLeuLysAenValSerPro 710
QY 3022 TTCTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGATCAA 3081
Db 711 PheValLysLeuGlnGlyValTyrrSerGluGlnArgLysPheSerGluGlyLeuArg 730
QY 3082 GCTCGGCGCATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTT 3141
Db 731 ArgCysLeuPheSerSerThrTyrrLeuAlaAenLeuAlaLeuProLeuGlyIleLysIle 750
QY 3142 GATCGATGTTCTAGTACACATCTTAATAATATAGCTTTTATGGCGCTTATATCTGTGAT 3201
Db 751 HisGlyValCysProArgGluLeuPheAlaTyrrAaspLeuSerAlaMetTyrrValHisAasp 770
QY 3202 GCTTATCGCACCATCTCTGTTACTGACAGCAACG---CTCCTATCCCATCAAGACATGG 3258
Db 771 ValPheArgileAaspProGluThrMetThrLeuPheLeuileGlyGlyLeuAlaProTrp 790
QY 3259 ACAACAGATGCTTTCATTTTAGCAAGACATGGAGTTGTTGTTAGAGGATCTATGATGCT 3318
Db 791 ThrThrHisAlaAenAenLeuAlaThrLysAlaileValValGlnGlySerGlyArgPhe 810
QY 3319 TCTTAACAAGTATATAGAAGTATATATGGCCATGAGAGATATAGATATCGAGATGCTTCT 3378
Db 811 AlaValArgSerAenileGluValPheAlaGluGlyAenCysGluLeuArgSerSerSer 830
QY 3379 CGAGCTTATGTTTGGAGTGGAGGAGTAGATCGGTTTC 3417
Db 831 HisSerTyrrAenTyrrAaspPheGlyAlaLysileHisPhe 843
```

Search completed: May 13, 2006, 10:31:16  
Job time : 1252.5 sec

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: May 13, 2006, 09:33:09 ; Search time 123.2 Seconds  
(without alignments)  
4745.077 Million cell updates/sec

Title: US-10-701-844-1  
Perfect score: 7883  
Sequence: 1 gggcaaaactctcccccgcg.....gcattcgaagaatttc 4435

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB/spool/US10701844/runat\_12052006\_165417\_26223/app\_query.fasta\_1  
-DB=A.Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNIT5=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02p  
-USER=US10701844 @CGN\_1\_1\_348 @runat\_12052006\_165417\_26223 -NCPU=6 -ICPU=3  
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq\_21.\*  
1: Genesecp1980s.\*  
2: Genesecp1990s.\*  
3: Genesecp2000s.\*  
4: Genesecp2001s.\*  
5: Genesecp2002s.\*  
6: Genesecp2003as.\*  
7: Genesecp2003bs.\*  
8: Genesecp2004s.\*  
9: Genesecp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	66.8	1012	2	Aay16735 C. tracho
2	5131.5	65.1	1013	5	Abg91021 Chlamydia
3	5131.5	65.1	1013	7	Add43798 Chlamydia
4	5131.5	65.1	1013	7	Add43722 Chlamydia
5	5131.5	65.1	1013	7	Add43718 Chlamydia
6	5131.5	65.1	1013	7	Add43708 Chlamydia
7	5131.5	65.1	1013	7	Add43714 Chlamydia
8	5131.5	65.1	1013	7	Add43786 Chlamydia
9	5131.5	65.1	1013	7	Add43788 Chlamydia

10	5131.5	65.1	1013	7	ADD43710	Chlamydia
11	5131.5	65.1	1013	7	ADD43706	Chlamydia
12	5131.5	65.1	1013	7	ADW29027	C. trachom
13	5131.5	65.1	1013	9	AEa19078	Chlamydia
14	5131.5	65.1	1013	9	AEa19002	Chlamydia
15	5125.5	65.0	1013	9	AEa19080	Chlamydia
16	5123.5	65.0	1013	9	AEa19006	Chlamydia
17	5123.5	65.0	1013	9	AEa19090	Chlamydia
18	5123.5	65.0	1013	9	AEa19010	Chlamydia
19	5121.5	65.0	1013	9	AEa19000	Chlamydia
20	5121.5	65.0	1013	9	AEa18998	Chlamydia
21	5112.5	64.9	1013	9	AEa19014	Chlamydia
22	5106.5	64.8	1013	2	AY16737	C. tracho
23	5106.5	64.8	1013	2	AY16738	C. tracho
24	5090	64.6	1006	4	AAG83207	Protein e
25	5090	64.6	1006	5	AB94178	Chlamydia
26	5084	64.5	982	3	AB13633	C. tracho
27	5084	64.5	982	3	AAG83201	Protein e
28	5084	64.5	982	5	AB94172	Chlamydia
29	5083	64.5	1006	3	AB13639	C. tracho
30	3336.5	42.3	670	7	ADD42756	Chlamydia
31	2606	33.1	524	2	AY37238	Chlamydia
32	2547	32.3	505	2	AY16739	C. tracho
33	2350	29.8	458	2	AY16751	Chlamydia
34	1735	22.0	325	2	AY16752	Chlamydia
35	1715	21.8	708	2	AY37234	Chlamydia
36	1590.5	20.2	631	4	AAG83274	Chlamydia
37	1590.5	20.2	631	5	AB94245	Chlamydia
38	1589	20.2	1016	4	AU38901	C. tracho
39	1589	20.2	1016	5	AB91037	Chlamydia
40	1589	20.2	1016	7	ADD42682	Chlamydia
41	1589	20.2	1016	7	ADD43800	Chlamydia
42	1589	20.2	1016	9	ADW29028	C. trachom
43	1589	20.2	1016	9	AEa19092	Chlamydia
44	1456	18.5	664	7	ADD42755	Chlamydia
45	1365.5	17.3	973	5	AB90527	Chlamydia

## ALIGNMENTS

RESULT 1  
AAY16735  
ID AAY16735 standard; protein; 1012 AA.  
XX  
AC AAY16735;  
XX

DT 21-JUL-1999 (first entry)  
XX

DE C. trachomatis LGV L2 HMW protein.  
XX

KW Chlamydia; high molecular weight protein; HMW protein; urethritis;  
KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;  
KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;  
KW salpingitis; tubal occlusion; infertility; cervical cancer;  
KW arteriosclerosis; atherosclerosis.

OS Chlamydia trachomatis.  
XX

PN WO9917741-A1.  
XX

PD 15-APR-1999.  
XX

PP 01-OCT-1998; 98WO-US020737.  
XX

PR 02-OCT-1997; 97US-00942596.  
XX

PA (ANTE-) ANTEX BIOLOGICS INC.  
XX

PI Jackson JW, Pace JL;  
XX

DR WPI; 1999-287659/24.  
XX

DR N-PSDB; AAX60539.  
XX



PT New Chlamydia protein useful for treating conjunctivitis, urethritis and  
 PT cervical cancer.

XX Claim 4; Fig 3; 141pp; English.

XX The invention relates to an isolated Chlamydia species high molecular  
 CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as  
 CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can  
 CC be used for preventing, treating or ameliorating a disorder related to  
 CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,  
 CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,  
 CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical  
 CC cancer, infertility, arteriosclerosis and atherosclerosis. The products  
 CC can also be used for detection and diagnosis. The present sequence  
 CC represents a C. trachomatis HMW protein

XX Sequence 1012 AA;

Alignment Scores:

Pred. No.: 0 Length: 1012  
 Score: 5267.00 Matches: 1012  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 66.8% Indels: 0  
 DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x AAY16735 (1-1012)

QY 382 ATGCAAGCTCTTTCATAGTCTCTTCAATGATTCTAGCTATTCTTGCTGCTCT 441  
 DB 1 MetGlnThrSerPheHisLeuPheLeuSerMetIleLeuAlaIleSerCysSer 20  
 QY 442 TTAATGGGGGATATGCGACAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501  
 DB 21 LeuAsnGlyGlyTyAlaAlaGluIleMetValProGlnGlyIleTyAspGlyGlu 40  
 QY 502 ACGTTAAGTATCATTTCCCTATATCTGTTATAGAGATCCGAGTGGGACTCTGTTTTT 561  
 DB 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60  
 QY 562 TCTGACGAGAGTTAAACATTAATAAATCTTGCAATCTTATGACGCTTTCGCTTTAAGT 621  
 DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
 QY 622 TGTGTTGGAACTTATTAGGAGTTTACTGTGTTTGGGAGAGGACACTCGTTGACTTTC 681  
 DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 QY 682 GAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741  
 DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120  
 QY 742 TTTACTATTGAGGGTTTAAAGAAATTCCTTTTCCAATTGCAATTCATTACTTCGCGTA 801  
 DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
 QY 802 CTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACACATCTACACCGTCT 861  
 DB 141 LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrThrSerThrProSer 160  
 QY 862 AATGTAATATTATTTCTAAACAGATCTTTGTTTACTCAATTAATGAGAGTTCTCATTC 921  
 DB 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
 QY 922 TATAGTAATTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981  
 DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
 QY 982 GGAATTAGCAAGCTTTGTCTTCCAAAGAAATATCTGCTCAAGCTGATGGGGAGCTGT 1041  
 DB 201 GlyIleSerLysLeuCysValPheGlnGlnLeuThrAlaGlnAlaAspGlyAlaCys 220  
 QY 1042 CAAGTAGTCACCAAGTTCTCTGCTATGGCTAAACGAGGCTCTATTGCGCTTTGTAGCGAAT 1101

DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240  
 QY 1102 GTTCAGAGTAAGAGGGGAGGATGCTGCTCTTTCAGGATGGGACAGCGAGTGTCA 1161  
 DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
 QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATATCTCGCGTAGAGTTT 1221  
 DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
 QY 1222 GATCGGACGTAGCCCGAGTAGGAGGAGGATTTACTCTCTACGGGAACGTGCTTCTCTG 1281  
 DB 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300  
 QY 1282 AATAATGGAAAAACCTTTGTTCTCAACAATGTTCTCTCTCTCTTATCATTTGCTGCTAAG 1341  
 DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyIleAlaAlaLys 320  
 QY 1342 CAACCAACAGTGCACAGGCTTCTTAATACGAGTAAATTAACGAGATGGAGGAGCTATC 1401  
 DB 321 GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyGlyAspGlyGlyAlaIle 340  
 QY 1402 TTCTGTAGATGTCGCGCAGCAGGATCCCAATTAACCTCTGATCAGTTTCTCTTGTATGATCGA 1461  
 DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly 360  
 QY 1462 CAGCGAGTAGTTTCTTTTAGTAGCAATGCTAGCTCTGCGAAAGGGGAGCTATTATATGCC 1521  
 DB 361 GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyAla 380  
 QY 1522 AAAAGCTCTCGTGTGCTAATCTGTCGCTCTCAATTTTAAAGAAATATCGCTAATGAT 1581  
 DB 381 LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp 400  
 QY 1582 GGTGCGAGCTTTTATAGGAGATCTGGAGACTCTGAGCTCAGTTTATCTGCTGATATGAGAT 1641  
 DB 401 GlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyroGlyAsp 420  
 QY 1642 ATTATTTTCGATGGAAATCTTAAAGAACACGCCAAAGAGATGCTGCCGATGTTAATGCGC 1701  
 DB 421 IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly 440  
 QY 1702 GTAATGTCCTCACAAGCCATTTTCGATGGATCGGAGGAGGAAATTAACGACATTAGA 1761  
 DB 441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIleThrThrLeuArg 460  
 QY 1762 GCTAAGCAGCGCATCAGATTCCTTTAATGATCCCATCGAGATGGCAACGGAATAAC 1821  
 DB 461 AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn 480  
 QY 1822 CAGCAGCGCAGTCTTCCAAACTCTTAAAAATTAACGATGGTGAAGGATACACAGGGAT 1881  
 DB 481 GlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGlyTyThrGlyAsp 500  
 QY 1882 ATTGTTTTGCTAATGGAAGCAGTACTTTGTACCAAAATGTTTACGATGAGCAGCAAGAGG 1941  
 DB 501 IleValPheAlaAsnGlySerSerThrLeuTyGlnAsnValThrIleGluGlnGlyArg 520  
 QY 1942 ATTGTTCTTCTGAAAGGCAAAATATCAGTGAATCTCTAAGTCACAGCAGGTGGAGT 2001  
 DB 521 IleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySer 540  
 QY 2002 CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACACAGCCT 2061  
 DB 541 LeuTyMetGluAlaGlySerThrTrpAspPheValThrProGlnProGlnGlnPro 560  
 QY 2062 CTGCGCTAATCAGTTGATCAGCTTTTCCAAATCTGCAATTTGCTCTTCTTTCTTTGTTA 2121  
 DB 561 ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeu 580  
 QY 2122 GCMAACATGAGTTACGAATCCTCCTACCAATCCTCCAGCGCAAGATTTCTCATCTGCA 2181

Db 581 AlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisProAla 600  
QY 2182 GTCATTGGTAGCACAACCTGCTGTTCTGTACAAATAGTGGCCCTATCTTTTTCAGGAT 2241  
Db 601 ValIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGluAsp 620  
QY 2242 TTGGATGATACAGCTTATGATAGGTATGATCGCTAGGTTCCTAAATCAAAAATCAATGTC 2301  
Db 621 LeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnIleAsnVal 640  
QY 2302 CTGAATATACAGTTAGGGACTAAGCCCCAGCTAAATGCCCCATCAGATTGTAGGG 2361  
Db 641 LeuIleLeuGlnLeuGlyThrIleYsProProAlaAsnAlaProSerAspLeuThrLeuGly 660  
QY 2362 AATGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCGGATCCTAATACA 2421  
Db 661 AsnGluMetProIleYsTyrGlyThrGlnGlySerTrpIleLeuAlaTrpAspProAsnThr 680  
QY 2422 GCAAAATAATGGTCTTATCTCTGAAAGCTACATGAGCTAAAGCTGGGTATATATCTGGG 2481  
Db 681 AlaAsnAsnGlyProTyrThrLeuIleYsAlaThrTrpThrIleYsThrGlyTyrAsnProGly 700  
QY 2482 CCTGAGCGAGTAGCTTCTTTGGTCCAAATAGTTATGGGATCCATTTTATAGATACGA 2541  
Db 701 ProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIleArg 720  
QY 2542 TCTGGCATTCAGCAATTCAGCAAGTGGATGGCGCTCTTATTTGTCGAGATTATGG 2601  
Db 721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeuTrp 740  
QY 2602 GTTCTGGAGTTTCGAATTTCTTCTATCATGCGGATGCTTTAGTTCAGGAGATATCGG 2661  
Db 741 ValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArg 760  
QY 2662 TATATTAGTGGGGTTATCTCTTAGGAGCAACTCTTACTTTGGATCATCGATGTTTGGT 2721  
Db 761 TyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPheGly 780  
QY 2722 CTAGCATTTACGAAGTATTTGTAGATCTAAAGATTATGATGTGTCTGTTCCATCAT 2781  
Db 781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsnHis 800  
QY 2782 CATGTTGTCATAGGATCCGTTTATCTATCTACCAACAAGCTTATGTGATCTCTATTGG 2841  
Db 801 HisAlaCysIleGlySerValTyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeu 820  
QY 2842 TTCGAGATGCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCA 2901  
Db 821 PheGlyAspAlaPheIleAlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSer 840  
QY 2902 TATACATTTGACAGAGAGCGATGTTCTGTTGGGATAATACCTGCTGCTGGCTGAGAGATT 2961  
Db 841 TyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysAlaLeuAlaGlyGluIle 860  
QY 2962 GGAAGCGGATACCGATTGATGTTACTCCATCTCAAGCTCTATTGATGATGTTGCGTCT 3021  
Db 861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuIleAsnGlnLeuLeuArgPro 880  
QY 3022 TTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGATCAA 3081  
Db 881 PheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAspGln 900  
QY 3082 GCTCGGCGATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGTGGAGTGAAGTTT 3141  
Db 901 AlaArgAlaPheLysSerGlyHisLeuLeuAsnLeuSerValProValGlyValLysPhe 920  
QY 3142 GATCGATGTTCTAGTACACATCTCTAATATAGCTTTTATGGCGCTTATATCTGTGAT 3201  
Db 921 AspArgCysSerSerThrHisProAsnLysTyrSerPheMetAlaAlaTyrIleCysAsp 940  
QY 3202 GCTTATCGCACCATCTCTGTTACTGACAGCAACGCTCTCTATCCCATCAAGAGACATGGACA 3261  
Db 941 AlaTyrArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThr 960

QY 3262 ACAGATGCTTTTCATTATTAGCAAGCATGGAGTTGTGGTTAGAGGATCTATGTATGCTTCT 3321  
Db 961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyrAlaSer 980  
QY 3322 CTAACAAGTAATATAGAAAGTATATGCGCATATGCGCAATATGAGATATGAGATGCTTCTCGA 3381  
Db 981 LeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGlyTyrArgAspAlaSerArg 1000  
QY 3382 GCGTATGCTTTTCAGTGCAGGAGTAGATGCTCGGTTTC 3417  
Db 1001 GlyTyrGlyLeuSerAlaGlySerArgValArgPhe 1012  
RESULT 2  
ABG91021  
ID ABG91021 standard; protein; 1013 AA.  
XX  
AC ABG91021;  
XX  
DT 29-NOV-2002 (first entry)  
XX  
DE Chlamydia trachomatis outer membrane protein G protein.  
XX  
KW Gram-negative bacterial bleb; PorB; outer membrane protein;  
KW Chlamydia trachomatis infection; Chlamydia pneumoniae infection;  
KW protective antigen; antibacterial; vaccine; gene; ds.  
XX  
OS Chlamydia trachomatis.  
XX  
PN WO200262380-A2.  
XX  
PD 15-AUG-2002.  
XX  
PF 08-FEB-2002; 2002WO-EP001356.  
XX  
PR 08-FEB-2001; 2001GB-00003169.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
PI Berthet FJ, Lobet Y, Poolman J, Verlant VGCL;  
XX  
DR WPI; 2002-657510/70.  
DR N-PSDB; ABS67342.  
XX  
XX Novel gram-negative bacterial bleb presenting on its surface PorB outer  
PT membrane protein from Chlamydia trachomatis or protective antigen from  
PT Chlamydia pneumoniae, useful for preventing Chlamydia infection.  
XX  
PS Disclosure; Page 10; 75pp; English.  
XX  
CC The present invention relates to a new gram-negative bacterial bleb  
CC presenting on its surface the PorB outer membrane protein from Chlamydia  
CC trachomatis, or a protective antigen from C. pneumoniae. The invention is  
CC useful for preventing C. trachomatis or C. pneumoniae infection in a  
CC host. The present nucleic acid sequence represents a Chlamydia  
CC trachomatis gene of the invention  
XX  
SQ Sequence 1013 AA;  
Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5131.50 Matches: 983  
Percent Similarity: 98.8% Conservativeness: 18  
Best Local Similarity: 97.0% Mismatches: 11  
Query Match: 65.1% Indels: 1  
DB: Gaps: 1  
US-10-701-844-1 (1-4435) x ABG91021 (1-1013)

QY 382 ATCAACAGCTTTTCCATAGATGTTCTTCTTCAATGATTCAGTATTATCTTCTGCTCTCT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20

QY 442 TTAATGGGGGGATATGACAGCAAAATCATGTTCTCTCAAGGAAATTTACGATGGGAG 501  
Db : : : : :  
QY 21 LeuSerGlyGlyTyAlaAlaGluLeuMetIleProGlnGlyIleTyAspGlyGlu 40  
QY 502 ACGTTAACTGATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTT 561  
Db : : : : :  
QY 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrValPhe 60  
QY 562 TCTCGAGGAGGTAAACATTTAAAAATCTTGACAAATCTATTCGACGCTTCCCTTAAGT 621  
Db : : : : :  
QY 61 SerAlaGlyGluLeuThrLeuLeuAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
QY 622 TGTTCGGAACTTATTAGGAGTTTACTGTTTGGGAGGAGGACACTCGTTGACTTTC 681  
Db : : : : :  
QY 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGAACATACGAGCTTCTACAAATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741  
Db : : : : :  
QY 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
QY 742 TTTACTATTGAGGGTTTTAAAGAATTATCTTTTCCAAATTCGAATTCATTACTTCCCGTA 801  
Db : : : : :  
QY 121 PheThrIleGluGlyPheGlyGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
QY 802 CTGCTGCTGCACGACTAATAAGGGTAGCCGAGCTCCGACGACAACTCTACACCGTCT 861  
Db : : : : :  
QY 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
QY 862 AATGGTACTATTATTCTTAAACAGACTCTTTTGTGTACTCAATAATGAGAAGTTCTCATTC 921  
Db : : : : :  
QY 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuLeuAsnAsnGluTyPheSerPhe 180  
QY 922 TATAGTAATTAGTCTCTCGAGATGGGGAGCTATAGATGCTTAAGACTTAACGGTTCAA 981  
Db : : : : :  
QY 181 TyrSerAsnLeuValSerGlyAspGlyAlaAlaLeuAspAlaLysSerLeuThrValGln 200  
QY 982 GGAATTAGCAAGCTTTGTGCTTCTCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db : : : : :  
QY 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
QY 1042 CAAGTAGTCACCAAGTTCTTGCTATGCTTAACAGAGCTCTTATTCCTTTGCTAGCGAAT 1101  
Db : : : : :  
QY 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
QY 1102 GTTCAGGAGTAAGAGGGGAGGATGCTGCTGCTCAGGATGGGCAGCGAGGTCA 1161  
Db : : : : :  
QY 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATCTCCGAGTACTCGGTAGGTTT 1221  
Db : : : : :  
QY 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
QY 1222 GATGGGAAGTAGCCGAGTAGGAGGATTTACTCTACGGAAAGCTGCTTTCTCTG 1281  
Db : : : : :  
QY 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300  
QY 1282 AATAATGGAACCTTGTTCTCAACATGTTGCTTCTCTGTTTACATTCCTCTAAG 1341  
Db : : : : :  
QY 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyIleAlaAlaGlu 320  
QY 1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATAATTACGGAGATGGAGGACTATC 1401  
Db : : : : :  
QY 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle 340  
QY 1402 TTCTGTGAAGATGGTGGCAA--GCAGGATCCAAATCTCGGATCAGTTTCTTTGAT 1458  
Db : : : : :  
QY 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
QY 1459 GGAGAGGGAGTAGTTTCTTTTAGTACAAATGATGCTGCTGGGAAAGGGGAGGACTATTAT 1518  
Db : : : : :  
QY 361 GlyGluGlyValValPhePheSerAsnValAlaAlaGlyLysGlyAlaIleTy 380  
QY 1519 GCCAAAAGCTCTCGGTTCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578

Db : : : : :  
QY 381 AlalysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
QY 1579 GATCGTCGAGCGATTTATTTAGGAAATCTGGAGACTCAGTTTATCTGCTGATATGGA 1638  
Db : : : : :  
QY 401 AspGlyGlyAlaIleTyLeuGlyLeuSerGlyGluLeuSerLeuSerAlaAspTyGly 420  
QY 1639 GATATTATTTTCGATGGGAATCTTTAAAGAACACCCAAAGAAATGCTGCCGATGTTAAT 1698  
Db : : : : :  
QY 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
QY 1699 GGCCTAACTGTCTCACAGCCATTTTCGATGGATCGGAGGCGAAATAACACCATTA 1758  
Db : : : : :  
QY 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
QY 1759 AGACTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACCGAAAT 1818  
Db : : : : :  
QY 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
QY 1819 AACAGCCAGCGCAGTCTTCCAACTTCTTAAAAATTAACGATGCTGAAGGATACACAGG 1878  
Db : : : : :  
QY 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyThrGly 500  
QY 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGGA 1938  
Db : : : : :  
QY 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly 520  
QY 1939 AGGATTGTTCTTCGTAAGGCGAAATTTACAGTAATCTCTAAGTCAGACAGGTGG 1998  
Db : : : : :  
QY 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
QY 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACCCACCAACAG 2058  
Db : : : : :  
QY 541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
QY 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCTCAATCTGCAATTTGCTCTTCTCTCTTG 2118  
Db : : : : :  
QY 561 ProProAlaAlaAsnGlnMetIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580  
QY 2119 TTAGCAACAATCAGATTACGAATCTCTCTACCAATCTCCAGCGCAAGATTCATCCT 2178  
Db : : : : :  
QY 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
QY 2179 CGAGTCATGTTAGCAACACTGCTGTTCTGTTACAAATTAGTGGCCTATCTTTTTCAG 2238  
Db : : : : :  
QY 601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
QY 2239 GATTTCGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298  
Db : : : : :  
QY 621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsp 640  
QY 2299 GTCTCGAAATACAGTTAGGACTTAAGCCCGCAGCTAATGCCCATCAGATTTGACTCTA 2358  
Db : : : : :  
QY 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
QY 2359 GGGATGAGATGCCCTAAGTATGGCTATCAAGGAGCTGGAGCTTGGGATCCTAAT 2418  
Db : : : : :  
QY 661 GlyAsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
QY 2419 ACAGCAAAATATGTCCTTATCTCTGAAAGCTACATGCACTATAAATCGGTATATCTCT 2478  
Db : : : : :  
QY 681 ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAsnPro 700  
QY 2479 GGGCTGAGCGAGTAGTCTCTTTTGGTCCAAATAGTTTATGGGATCCAAATTTAGATATA 2538  
Db : : : : :  
QY 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
QY 2539 CGATCTGGCATTTCAGCAATTCAGCAAGTCTGGATGGGCTCTTATTTGTCAGGATTA 2598  
Db : : : : :  
QY 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740  
QY 2599 TGGGTTTCTGAGATTTCGAATTTCTTCTATCATGACCGCGATGCTTTTAGTCAGGATAT 2658

Db 741 TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
 Qy 2659 CGGTATATTAGTGGGGTTATTCCTTAGAGCAAACTCCTACTTTGGATCATCGATGTTT 2718  
 Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerMetPhe 780  
 Qy 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTGCCAAT 2778  
 Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
 Qy 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTATACCCAAAGCTTTATGTGCATCTAT 2838  
 Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
 Qy 2839 TTGTTCCGAGATCGGTTTATCCGCTAGCTACGGTTTGGCAATCGCATATGAACC 2898  
 Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
 Qy 2899 TCATATACATTTGCAGAGGAGCGATGTTGCTGGGATAATAACTGTCTGCTGGAGAG 2958  
 Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
 Qy 2959 ATTGAGCGGGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGTAATGAGTGGCT 3018  
 Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
 Qy 3019 CTTTCTGTCAGCTGAGTTTCTTATGCGCATCATGATCTTTACAGAGGAGCGAT 3078  
 Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGlyAsp 900  
 Qy 3079 CAGCTCGGGCTTCAAGAGCGACATCTCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138  
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
 Qy 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGGCTTATPCTGT 3198  
 Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
 Qy 3199 GATGCTTATCGACCACTCTGCTGATGAGCAACGCTCTATCCATCCATCAGAGCATGG 3258  
 Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
 Qy 3259 ACAACAGATCGCTTTCATTTAGCAACATGAGGATGTTGTGTTAGAGATCTATGATGCT 3318  
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
 Qy 3319 TCTCTAACCAAGTAAATAGAGTATATGCCATGGAAGATATGAGTATCGAGATGCTTCT 3378  
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
 Qy 3379 CGAGGCTATGTTTGTAGTGAGGAGTAGAGTCCGGTTC 3417  
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013  
 RESULT 3  
 ID ADD43798  
 AC ADD43798 standard; protein; 1013 AA.  
 XX ADD43798;  
 DT 15-JAN-2004 (first entry)  
 XX Chlamydia trachomatis immunogenic protein, SEQ ID No 93.  
 DE Chlamydia trachomatis immunogenic protein, SEQ ID No 93.  
 KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
 OS gene therapy; antibacterial.  
 XX Chlamydia trachomatis.  
 OS Chlamydia trachomatis.  
 XX WO2003049762-A2.  
 PN 19-JUN-2003.  
 PD 19-JUN-2003.  
 XX

PF 12-DEC-2002; 2002WO-IB005761.  
 XX 12-DEC-2001; 2001GB-00029732.  
 PR 06-AUG-2002; 2002GB-00018233.  
 PR 14-AUG-2002; 2002GB-00018924.  
 XX (CHIR-) CHIRON SPA.  
 PA Grandi G, Ratti G;  
 XX WPI; 2003-532882/50.  
 DR N-PSDB; ADD43799.  
 XX New immunogenic composition having a protein or encoding nucleic acid,  
 PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
 PT infection.  
 PS Claim 6; SEQ ID NO 93; 164pp; English.  
 XX The invention relates to a novel immunogenic composition comprising a  
 CC protein or nucleic acid, and an adjuvant, where the protein or nucleic  
 CC acid comprises any of 131 fully defined amino acid or nucleotide  
 CC sequences given in the specification, or has 50% or greater sequence  
 CC identity to it, or their fragments. The protein and/or nucleic acid of  
 CC the immunogenic composition is useful in the manufacture of a medicament  
 CC for the treatment or prevention of infection due to Chlamydia  
 CC trachomatis. The infection is treated or prevented by the medicament  
 CC eliciting an immune response which is specific to a C. trachomatis  
 CC elementary body, or for neutralising C. trachomatis elementary bodies,  
 CC hence the immunogenic composition can be used in creating a vaccine. The  
 CC immunogenic compositions can also be used for the diagnosis of C.  
 CC trachomatis infection. The nucleic acids of the immunogenic compositions  
 CC can be used to treat disorders by gene therapy. The immunogenic  
 CC compositions have antibacterial activity. This sequence represents one of  
 CC the 131 C. trachomatis proteins with immunogenic properties of the  
 CC invention.  
 XX SQ Sequence 1013 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 1013  
 Score: 5131.50 Matches: 983  
 Percent Similarity: 98.8% Conservative: 18  
 Best Local Similarity: 97.0% Mismatches: 11  
 Query Match: 65.1% Indels: 1  
 DB: 7 Gaps: 1  
 US-10-701-844-1 (1-4435) x ADD43798 (1-1013)  
 Qy 382 ATGCAACCTCTTTCATTAAGTTCTTTCTTCAATGATCTAGCTTATTCTGCTCTCT 441  
 Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCys 20  
 Qy 442 TTAATATGGGGGGATATGACAGAAATCATGGTTCCTCAAGAAATTTACGATGGGAG 501  
 Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
 Qy 502 ACCTTAACCTGATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGACTCTGTTT 561  
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
 Qy 562 TCTGACGAGGAGTTAAACATTTAAATAATCTTGACAACTTATTGACGCTTTGCTTTAAGT 621  
 Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80  
 Qy 622 TGTGTTGGGAACCTTATTAGGAGCTTTTACTGTTTATAGGAGAGACACTCGTTGACTTTC 681  
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 Qy 682 GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAGTAATAGCGCTGCTGATGGACTG 741  
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120

Qy	742	TTTACTATTGAGGGTTTAAAGAATTATCTTTTCCAAATGTCGAATTCATTACTTGCCGTA	801
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
Qy	802	CTGCTGCTGCAACGACTAATAAGGTTAGCCAGACTCCGACGACACATCTACACCGTCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
Qy	862	AATGGTACTATTATTCTAAACAGACTTTTGTGTACTCAATATGAGAAGTTCTCATTC	921
Db	161	AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe	180
Qy	922	TATAGTAAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln	200
Qy	982	GGAATTAGCAAGCTTTGTGCTCTCCACAGAAAATACTGCTCAAGCTGATGGGAGCTTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
Qy	1042	CAAGTAGTACCAGTTTCTCTGCTTATGGCTAAACGAGGCTCTATTGCTCTTGTAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGlnAlaProIleAlaPheIleAlaAsn	240
Qy	1102	GTTCACGAGTAGAAGGGGAGGATGCTGCTGTTTCAGGATGGCAGCAGGAGTGTCA	1161
Db	241	ValAlaGlyValAlaGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlyValSer	260
Qy	1162	TCATCTACTTCAACAGAAATCCAGTAGTAGTAAAGTTTTCAGAAATACTCGGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
Qy	1222	GATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTCCCTG	1281
Db	281	AspGlyAsnValAlaAlaGValGlyGlyIleTySerTyGlyAsnValAlaPheLeu	300
Qy	1282	AATAATGGAAAAACCTTGTTTCTCAACAATGTGCTTCTCTGTTTACATTCGTGCTAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrlleAlaAlaGlu	320
Qy	1342	CAACCAACAGTGGACAGGCTTCTAATACGAGTAATTAATTACGAGATGAGGAGCTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle	340
Qy	1402	TTCTGTAAAGATGTTGCGCAA--GCAGGATCCAAATAACTCTGATCAGTTTCTCTTGAT	1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
Qy	1459	GGAGAGGAGTAGTATTTCTTTAGTAGCAATGACTGCTGGAAAGGGGAGCTATTAT	1518
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTy	380
Qy	1519	GCCAAAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT	1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
Qy	1579	GATGTTGAGAGCATTTATTTAGGAGATCTCGAGAGCTCAGTTTATCTGCTGATTATGGA	1638
Db	401	AspGlyGlyAlaIleTyLeuGlyLysSerGlyGluLeuSerLeuSerAlaAspTyGly	420
Qy	1639	GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAATGCTGCCGATGTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
Qy	1699	GGCGTAATGTGTCTCAACAGCCATTTTCGATGGGATCGGGAGGGAAAAATAACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu	460
Qy	1759	AGAGCTAAGCAGGCATCTCAGTTCTTTAATGATCCATCGAGATGGCAACCGGAAT	1818
Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
Qy	1819	AACCAACGACGCGCAGCTCTTCCAAAATCTCTAAAAAATTACGATGTGTGAAGGATACACAGG	1878

Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
Qy 2959 ATTGAGCGGATTACCATGTTGATTACTCCATCTAAGCTTATTGAATGAGTTGGT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
Qy 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy 3079 CAAGCTCGGCAATCAAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
Qy 3139 TTTCATCATGTTCTAGTACACATCTTAATAATAGCTTTATGCGGCTTATATCTGT 3198  
Db 921 PheAspArgCysSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
Qy 3199 GATGCTTATCGCACCATCTCTGTTACTGAGACAGCGCTCTATCCATCAAGAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACAACAGATGCTTTTCATTAGCAAGACATGAGTGTGGTTAGAGGATCTATGATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
Qy 3319 TCTCTAACAGCTAATATAGTATATGCGCATGATGCGGATGATGATGATGATGATGAT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
Qy 3379 CGAGCTATGTTTTCAGTGCAGGAAGTAGAGTCCCGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

## RESULT 4

ID ADD43722 standard; protein; 1013 AA.  
XX  
AC ADD43722;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 17.  
XX  
KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
KW gene therapy; antibacterial.  
XX  
OS Chlamydia trachomatis.  
XX  
PN WO2003049762-A2.  
XX  
PD 19-JUN-2003.  
XX  
PF 12-DEC-2002; 2002WO-IB005761.  
XX  
PR 12-DEC-2001; 2001GB-00029732.  
PR 06-AUG-2002; 2002GB-00018233.  
PR 14-AUG-2002; 2002GB-00018924.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Grandi G, Ratti G;  
XX  
XX WPI; 2003-532882/50.  
DR N-PSDB; ADD43723.  
XX  
PT New immunogenic composition having a protein or encoding nucleic acid,  
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
PT infection.  
XX  
PS Claim 6; SEQ ID NO 17; 164pp; English.  
XX  
XX The invention relates to a novel immunogenic composition comprising a

CC protein or nucleic acid, and an adjuvant, where the protein or nucleic  
CC acid comprises any of 131 fully defined amino acid or nucleotide  
CC sequences given in the specification, or has 50% or greater sequence  
CC identity to it, or their fragments. The protein and/or nucleic acid of  
CC the immunogenic composition is useful in the manufacture of a medicament  
CC for the treatment or prevention of infection due to Chlamydia  
CC trachomatis. The infection is treated or prevented by a C. trachomatis  
CC eliciting an immune response which is specific to a C. trachomatis  
CC elementary body, or for neutralising C. trachomatis elementary bodies,  
CC hence the immunogenic composition can be used in creating a vaccine. The  
CC immunogenic compositions can also be used for the diagnosis of C.  
CC trachomatis infection. The nucleic acids of the immunogenic compositions  
CC can be used to treat disorders by gene therapy. The immunogenic  
CC compositions have antibacterial activity. This sequence represents one of  
CC the 131 C. trachomatis proteins with immunogenic properties of the  
CC invention.

SQ Sequence 1013 AA;

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5131.50 Matches: 983  
Percent Similarity: 98.8% Conservative: 18  
Best Local Similarity: 97.0% Mismatches: 11  
Query Match: 65.1% Indels: 1  
DB: 7 Gaps: 1

US-10-701-844-1 (1-4435) x ADD43722 (1-1013)

Qy 382 ATGCAACGCTCTTCCATTAAGTTCTTCTTCAATGATCTAGCTTATTTCTGCTCTT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20  
Qy 442 TTAATGCGGGGATATGACAGCAATCATGGTTCTCTCAAGGAATTTACGATGGGAG 501  
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyLysPheGlyGlu 40  
Qy 502 ACGTTAACTGTATCATTTCCCTATATCTGTATAGGAGATCCGAGTGGGACTACTGTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTGAGGAGAGTTAAACATTAATAAATCTTGAATTTCTATTTGAGCTTTGGCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
Qy 622 TGTTCGGAACCTTATTAGGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy 682 GAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTCTGATGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
Qy 742 TTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCAATGCAATTCATTTACGCGTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnAsnSerLeuLeuAlaVal 140  
Qy 802 CTGCTGCTGCAACGACTAATTAAGGTAGCCAGACTCCGACGACACATCTACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
Qy 862 AATGGTACTATTATTCTAAACACAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
Qy 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGGAGCTTCT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220

1042 CAAGTAGTCACCAAGTTCTCTGCTATGCTAAACGAGGCTCCTATTGCTTTGTAGCGAAT 1101  
Db |||||  
221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
Qy |||||  
1102 GTTTCAGAGTAAGAGGGGAGGAGTTCGCTGCTTTTCAGAGTGGCGAGGAGGTGCA 1161  
Qy |||||  
241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
Db |||||  
1162 TCATCTACTTCACAGAGATCCAGTAGTAGTTTCCGAAATATCTGCGGTAGAGTTT 1221  
Qy |||||  
261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
Db |||||  
1222 GATGGGACGTAGCCGAGTAGGAGGAGGATTACTCTCTACGGGAAAGTTTCTCTCTG 1281  
Qy |||||  
281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
Db |||||  
1282 AATAATGGAAAACTTTGTTCTCAACAATGTTCTCTCTGTTTACATTGCTGCTAAG 1341  
Qy |||||  
301 AsnAsnGlyIleThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320  
Db |||||  
1342 CAACCAACAGTGCACAGGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401  
Qy |||||  
321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
Db |||||  
1402 TTCTGTAAAGATGGTGGCAA---GCAGGATCCCAATTAATCTGGATCAGTTTCTCTGAT 1458  
Qy |||||  
341 PheCysIleAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
Db |||||  
1459 GGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGTGGAAAGGGGAGCTATTAT 1518  
Qy |||||  
361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIleGlyAlaIleTyr 380  
Db |||||  
1519 GCCAAAAGCTCTGGTTGCTAACTGTGGCTGTACAAATTTTAAAGCAATATCGCTAAT 1578  
Qy |||||  
381 AlaIleIleLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
Db |||||  
1579 GATGTCGAGCGATTTATTAGGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
Qy |||||  
401 AspGlyGlyAlaIleTyrIleGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420  
Db |||||  
1639 GATATTATTTTCGAGTGGAACTTTAAAGAACAGCAAGAAAGTCTGCCGATTTAAT 1698  
Qy |||||  
421 AspIleIlePheAspGlyAsnLeuIleYsArgThrAlaIleGluAsnAlaAlaAspValAsn 440  
Db |||||  
1699 GCGCTAATCTCTCAACGCTTTCGATGGATCGGAGGAGAAATAACGACATTA 1758  
Qy |||||  
441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyIleThrThrLeu 460  
Db |||||  
1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818  
Qy |||||  
461 ArgAlaIleAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
Db |||||  
1819 AACCAAGCCAGCGAGCTTTCCAAACTTTAAATAATAACGATGGTGAGGATACACAGG 1878  
Qy |||||  
481 AsnGlnProAlaGlnSerSerGluProLeuIleAsnAspGlyGlyIleThrGly 500  
Db |||||  
1879 GATATTGTTTTCGTAATGAGACGACTTTGTACCAAAATGTTACGATAGCAAGCA 1938  
Qy |||||  
501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
Db |||||  
1939 AGGATGTTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGG 1998  
Qy |||||  
521 ArgIleValLeuArgGluIleAlaIleLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
Db |||||  
1999 AGTCTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACACCAACAG 2058  
Qy |||||  
541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
Db |||||  
2059 CTTCTGCGCTTAATCAGTTGATCAGCTTCCCAATCTGATTTGCTCTTTCTCTTTG 2118  
Qy |||||  
561 ProProAlaAlaAsnGlnIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
Db |||||  
2119 TTAGCAAAACAAATGACGTACGAATCTCTCTACCAATCTCTCCAGCGCAAGATTCTCATCT 2178

581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
Qy |||||  
2179 GCAGTCAATCGTAGCAAACTGCTGTTCTGTGTACAAATAGTGGCCCTATCTTTTTCAG 2238  
Db |||||  
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
Qy |||||  
2239 GATTTCGATGATACAGCTTATGATAGTATGATGGCTAGGTTCTTAATCAAAAAATCAAT 2298  
Db |||||  
621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnIleAsp 640  
Qy |||||  
2299 GTCTCTGAAATACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTTCACTCTA 2358  
Db |||||  
641 ValLeuIleLeuGlnIleGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
Qy |||||  
2359 GGGAAATCAGATCGCTAAGTATGGCTATCAAGAGAGCTGGAAGCTTGGCTGGATCTTAAT 2418  
Db |||||  
661 GlyAsnGluMetProIleTyrGlyIleGlySerTrpIleLeuAlaIleTrpAspProAsn 680  
Qy |||||  
2419 ACAGCAAAATATGCTCTTATCTCTCAAAAGCTACATGGACTAAACCTGGGTATAATCT 2478  
Db |||||  
681 ThrAlaAsnAsnGlyProTyrThrLeuIleAlaThrTrpThrIleThrGlyTyrAsnPro 700  
Qy |||||  
2479 GGGCTGAGCGAGTAGCTTCTTGGTTCCAAATAGTTTATGGGATCCATTTTATAGATATA 2538  
Db |||||  
701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
Qy |||||  
2539 CGATCTCGCAATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTTGTCGAGGATTA 2598  
Db |||||  
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
Qy |||||  
2599 TGGGTTTCTGAGGATTCGAATTTCTTCATCATGACCGCGATGCTTTAGTTCAGGATAT 2658  
Db |||||  
741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
Qy |||||  
2659 CGGTATATTAGTGGGTTTATCTCTAGGACAACTCTACTTTGGATCATCGATGTTT 2718  
Db |||||  
761 ArgTyrIleSerGlyIleTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
Qy |||||  
2719 GGTCTAGCATTTACCGAAGTATTTGTAGATCTAAAGATATTAGTAGTGTGCTGTTCCAAAT 2778  
Db |||||  
781 GlyLeuAlaPheThrGluValPheGlyArgSerIleAspTyrValValCysArgSerAsn 800  
Qy |||||  
2779 CATCATGCTTCATAGGATCCGTTTATCTATCTATCTACCCCAAGCTTTATGAGGATCTCTAT 2838  
Db |||||  
801 HisHisAlaCysIleGlySerValTyrLeuSerThrIleGlnAlaLeuCysGlySerTyr 820  
Qy |||||  
2839 TTGTTCCGAGATCGGTTTATCCGTCGTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898  
Db |||||  
821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetIleThr 840  
Qy |||||  
2899 TCATATACATTTGACAGAGAGAGCGATGTTGTTGGGATTAATACTGTCTGCTGGAGAG 2958  
Db |||||  
841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
Qy |||||  
2959 ATTTGAGCGGATTAACGATTTGATCTCATCTAAGCTCTATTGATGATGAGTTCGCT 3018  
Db |||||  
861 IleGlyValGlyLeuProIleValIleThrProSerIleLeuTyrIleAsnGluLeuArg 880  
Qy |||||  
3019 CTTTCTGTCGCAAGCTGAGTTTCTTATGCGGATCATGATCTTTTACAGGAAGCGAT 3078  
Db |||||  
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy |||||  
3079 CAAGCTCGGCAATCAAGAGCGGACATCTCTCAAAATCTATCAGTTCTCTGTTGAGTGAAG 3138  
Db |||||  
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLeu 920  
Qy |||||  
3139 TTTGATTCGATGTTCTAGTACATCTCTAATAATATAGCTTTTATGGCGCTTATCTGT 3198  
Db |||||  
921 PheAspArgCysSerSerThrHisProAsnIleTyrSerPheMetGlyAlaTyrIleCys 940  
Qy |||||  
3199 GATGCTTATCGCAACCATCTCTGTCGTCGAGACAAGCTCTCTATCCCATCAAGAGACATGG 3258



Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTTTTCATTAGACAGACATGAGGTTGGTTAGAGGATCTATGATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
QY 3319 TCTCTAACCAAGTAATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
QY 3379 CGAGGCTATGTTGAGTGCAGGAGTATAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 5

ADD43718  
ID ADD43718 standard; protein; 1013 AA.

XX ADD43718;

XX 15-JAN-2004 (first entry)

XX Chlamydia trachomatis immunogenic protein, SEQ ID NO 13.

XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;

XX gene therapy; antibacterial.

XX Chlamydia trachomatis.

XX WO2003049762-A2.

XX 19-JUN-2003.

XX 12-DEC-2002; 2002WO-1B005761.

XX 12-DEC-2001; 2001GB-00029732.

XX 06-AUG-2001; 2002GB-00018233.

XX 14-AUG-2002; 2002GB-00018924.

XX (CHIR-) CHIRON SPA.

XX Grandi G, Ratti G;

XX WPI; 2003-532882/50.

XX N-PSDB; ADD43719.

XX New immunogenic composition having a protein or encoding nucleic acid,  
XX useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
XX infection.

XX Claim 6; SEQ ID NO 13; 164pp; English.

XX The invention relates to a novel immunogenic composition comprising a  
XX protein or nucleic acid, and an adjuvant, where the protein or nucleic  
XX acid comprises any of 131 fully defined amino acid or nucleotide  
XX sequences given in the specification, or has 50% or greater sequence  
XX identity to it, or their fragments. The protein and/or nucleic acid of  
XX the immunogenic composition is useful in the manufacture of a medicament  
XX for the treatment or prevention of infection due to Chlamydia  
XX trachomatis. The infection is treated or prevented by the medicament  
XX eliciting an immune response which is specific to a C. trachomatis  
XX elementary body, or for neutralising C. trachomatis elementary bodies,  
XX hence the immunogenic composition can be used in creating a vaccine. The  
XX immunogenic compositions can also be used for the diagnosis of C.  
XX trachomatis infection. The nucleic acids of the immunogenic compositions  
XX can be used to treat disorders by gene therapy. The immunogenic  
XX compositions have antibacterial activity. This sequence represents one of  
XX the 131 C. trachomatis proteins with immunogenic properties of the  
XX invention.

XX Sequence 1013 AA;

XX Alignment Scores:

Pred. No.: 0 Length: 1013  
Score: 5131.50 Matches: 983  
Percent Similarity: 98.8% Conservative: 18  
Best Local Similarity: 97.0% Mismatches: 11  
Query Match: 65.1% Indels: 1  
DB: 7 Gaps: 1  
US-10-701-844-1 (1-4435) x ADD43718 (1-1013)  
QY 382 ATGCAAAAGCTCTTTCCATTAAGTCTCTTCTTCAATGATTCTAGCTTATTCTTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20  
QY 442 TTAATATGGGGGGATATGACAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501  
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
QY 502 ACGTTAACTGTATCATTTCCCTATATCTGTATAGAGATCCGAGTGGGACTACTGTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGCAGGAGACTTAACATTTAAATAATCTTGCAATTTCTATTGACGCTTTGCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
QY 622 TGTTTTGGGAACCTTATTAGGGAGTTTTACTGTGTTTTAGGGAGAGGACACTCGTTGACTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGAACATACGACTTCTACAAATGGGGCAGCTCTAAGTAATAGCCCTGCTGATGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTTCCCAATTCATTTACTTCCGCTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
QY 802 CTGCTCTGCTGCAACGACTAATAAGGCTAGCCAGACTCCGACGACAAACATCTACACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
QY 862 AATGGTACTATTATTCTAAACAGATCTTTTGTACTCAATATGAAAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnGluLysPheSerPhe 180  
QY 922 TATAGTAATTTACTCTCTCGAGATGGGGAGCTATAGATGCTAAGAGCTTAAACGGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
QY 982 GGAATTAGCAAGCTTTGTGCTTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
QY 1042 CAAGTAGTACCAAGTTTCTCTGCTATGGCTTAACGAGGCTCTATTGCTTTGATGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
QY 1102 GTTGCAGGAGTATAGGGGGAGGAGTTGCTGCTGTTCCAGATGCGGACGAGGAGTCTCA 1161  
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTTCCAGAATACTCGCGGTAGAGTTT 1221  
Db 261 SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
QY 1222 GATGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTTACGGAAACGTTGCTTCTCTG 1281  
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
QY 1282 AATAATGGAAAAACCTTGTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTTGCTGCTAAG 1341  
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320

```
QY 1342 CAACCAACAGTGGACGGCTCTTAATACGAGTAATAATACGGAGATGGAGGACTATC 1401
Db |||||:::|||||
QY 1402 TTCTGTAGAATGGTGGCCAA---CGAGGATCCAAATACTCTGGATCAGTTTCCTTTGAT 1458
Db |||||PheCysLeuAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 GGAGAGGAGTAGTGTCTTTAGTAGCAATGTAGCTGTGGAAAGGGGAGCTATTTAT 1518
Db |||||GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLeuGlyAlaAlaIleTyr 380
QY 1519 GCCAAAAGCTCTCGTTCTTAACGTGGCCCTGTACAAATTTTAAGCAATATCGCTAAT 1578
Db |||||AlaIleLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
QY 1579 GATGTCGAGCGCATTTATTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
Db |||||AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAATGTCGCCGATCTTAAT 1698
Db |||||AspIleIlePheAspGlyAsnLeuLeuysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GCGTAACTGTCTCTCACAGCCATTTTCGATGGGATCGGAGGAGAAATAACGCATTA 1758
Db |||||GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLeuIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACCGAAAT 1818
Db |||||ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACCCAGCCAGCCAGCTCTTCCAAACTCTTAAATAATTAACGATGGTGAAGGATACACAGG 1878
Db |||||AsnGlnProAlaGlnSerSerGlyProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
QY 1879 GATATTGTTTTCCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACGAAGA 1938
Db |||||AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATGTTCTTCGTGAAAGGCCAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
Db |||||ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCGTATATGAAGCTGGAGTACATGGGATTTTGTAACTTCCCAACACCACCAACAG 2058
Db |||||SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
QY 2059 CCTCTGCGCTAATCAGTTGATCAGCTTTTCCAAATCTGCATTTCTCTTCTTCTTTG 2118
Db |||||ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAAAACAATGCAAGTACGAATCTCTTACCAATCTCCAGCGCAAGATTTCTATCCT 2178
Db |||||LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 CGAGTCATTTGGTACACAACTCGTCTGTGTACAAATPAGTGGCCCTATCTTTTGGAG 2238
Db |||||AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTTGGATGATACAGCTTATCATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298
Db |||||AspLeuAspAspThrAlaTyrAspArgTyrAspTyrPheLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTCGAAATACAGTTAGGACTTAAGCCCCAGCTTAATGCCCACTCAGATTTGACTCTTA 2358
Db |||||ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GCGAATAGATCCCTAAGTATCGCTATCAAGAACTGGAAGCTGCGTGGGATCCTAAT 2418
Db |||||GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
QY 2419 ACAGCAATAATGGTCTTATCTCTGAAAGCTACATGGACTAAAACTGGGTATAATCCT 2478
```

```
Db |||||ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCTGAGCGAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTATAGATATA 2538
Db |||||GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGGCATTTACGCAATTCACCAAGTGTGATGGCGCTCTTATTTGTCGAGGATTA 2598
Db |||||ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGAGTTTCGAATTTCTTCTATCATCACCGCATGCTTTTAGTTCAGGATAT 2658
Db |||||TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCTTACGAGCAAACTCTACTTTGGATCATCATGATGTTT 2718
Db |||||ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGATGTCGTGCTTCCAAT 2778
Db |||||GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCCAAAGCTTTTATGTGGATCCTAT 2838
Db |||||HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTTCGAGATGCGTTTATCCGTCAGCTACCGGTTTGGGAATTCAGCATATGAAAAACC 2898
Db |||||LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGGAGGATGTCGTTGGGATATAACTGTCGTGCTGGAGAG 2958
Db |||||SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGATTTACCGATTTGTGATTACTCCATCTAAGCTCTATTTCAATGAGTTGCGT 3018
Db |||||IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CCTTTCGTGCAAGCTGAGTTTCTTTATGCCGATCATGAATCTTTTACAGAGAAAGCGCAT 3078
Db |||||ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGTCGGGATTTCAAGAGCGGACATCTCTTAATCTCATGTTCTCTGCTGGAGTGAAG 3138
Db |||||GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCCATGTTCTAGTACACATCTCTAATAAATATAGCTTTATGGCGCTTATATCTGT 3198
Db |||||PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGACCATCTCTGGTACTGAGCAACGCTCTCTATCCATCCATCAAGAGACATGG 3258
Db |||||AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGlnThrTrp 960
QY 3259 ACAACAGATGCTTTCATTTAGCAAGACATGGAGTTGTTGGTTAGGATCTATGATGCT 3318
Db |||||ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAAACAAGTAATATAGAGATATATGCCCATGGAAGATATGATATGAGATGCTTCT 3378
Db |||||SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGCTATGGTTTGGATGCGAGGAAGTAGAGTCGCGTTTC 3417
Db |||||ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 6
ADD43708
ID ADD43708 standard; protein; 1013 AA.
XX
AC ADD43708;
XX
```

DT 15-JAN-2004 (first entry)  
 XX Chlamydia trachomatis immunogenic protein, SEQ ID No 3.  
 DE immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
 XX gene therapy; antibacterial.  
 KW Chlamydia trachomatis.  
 XX WO2003049762-A2.  
 XX 19-JUN-2003.  
 XX 12-DEC-2002; 2002WO-1B005761.  
 XX 12-DEC-2001; 2001GB-00029732.  
 PR 06-AUG-2002; 2002GB-00018233.  
 PR 14-AUG-2002; 2002GB-00018924.  
 XX (CHIR-) CHIRON SPA.  
 PA Grandi G, Ratti G;  
 XX WPI: 2003-532882/50.  
 DR N-PSDB; ADD43709.  
 XX New immunogenic composition having a protein or encoding nucleic acid,  
 PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
 PT infection.  
 XX Claim 6; SEQ ID NO 3; 164pp; English.  
 XX The invention relates to a novel immunogenic composition comprising a  
 CC protein or nucleic acid, and an adjuvant, where the protein or nucleic  
 CC acid comprises any of 131 fully defined amino acid or nucleotide  
 CC sequences given in the specification, or has 50% or greater sequence  
 CC identity to it, or their fragments. The protein and/or nucleic acid of  
 CC the immunogenic composition is useful in the manufacture of a medicament  
 CC for the treatment or prevention of infection due to Chlamydia  
 CC trachomatis. The infection is treated or prevented by the medicament  
 CC eliciting an immune response which is specific to a C. trachomatis  
 CC elementary body, or for neutralising C. trachomatis elementary bodies,  
 CC hence the immunogenic composition can be used in creating a vaccine. The  
 CC immunogenic compositions can also be used for the diagnosis of C.  
 CC trachomatis infection. The nucleic acids of the immunogenic compositions  
 CC can be used to treat disorders by gene therapy. The immunogenic  
 CC compositions have antibacterial activity. This sequence represents one of  
 CC the 131 C. trachomatis proteins with immunogenic properties of the  
 CC invention.  
 XX  
 SQ Sequence 1013 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 1013  
 Score: 5131.50 Matches: 983  
 Percent Similarity: 98.8% Conservative: 18  
 Best Local Similarity: 97.0% Mismatches: 11  
 Query Match: 65.1% Indels: 1  
 DB: 7 Gaps: 1  
 US-10-701-844-1 (1-4435) x ADD43708 (1-1013)  
 QY 382 ATGCACACGCTTTCCATAAGTCTTTCTTCAATGATTCAGCTTATCTTGCTGCTCT 441  
 Db 1 MetGlnThrSerPheHisLeuPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20  
 QY 442 TTAATGGGGGGGATATGACGACCAATCATGTTCTTCAAGCAATTTACGATGGGAG 501  
 Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
 QY 502 ACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTT 561  
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrValPhe 60

QY 562 TCTGCAGGAGGTTAAACATTAATAAATCTTGACAAATCTATTGACAGCTTGGCTTTAAGT 621  
 Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
 QY 622 TGTGTTGGAACTTATTAGGAGTGTCTTACTGTTTGGAGGAGGAGGACACTCGTGTGACTTTC 681  
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 QY 682 GAGAACATACGACCTTCTACAAATGGGCGAGCTCTAAAGTAATAGCGCTGCTCATGACATG 741  
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
 QY 742 TTTTACTATTGAGGGTTTTAAAGAATTATCTCTTTTCCAATTGCAATTCAATTCTGCGGTA 801  
 Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
 QY 802 CTGCGCTGTCGAACGACTAATTAAGGGTAGCGAGCTCCGACGACCAACATCTACACGCTCT 861  
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
 QY 862 AATGCTACTATTATTCTAAACACAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTC 921  
 Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe 180  
 QY 922 TATAGTAATTTAGTCTCTCGAGATGGGCGAGCTATAGATGCTAAAGAGCTTAAACGGTCAA 981  
 Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
 QY 982 GGAATTAGCAAGCTTTGCTCTTCCAAAGAAATATCTGCTCAAGCTGATGGGAGCTTGT 1041  
 Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
 QY 1042 CAAAGTAGTCACCAAGTTCTCTGCTAGCTTAACGAGGCTCTATTGCTTTGAGCGAAT 1101  
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
 QY 1102 GTTGCAGGAGTAGTAAGGGGAGGAGTGTCTGCTGTTTCCAGGATGGGACGAGGAGGTGTC 1161  
 Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
 QY 1162 TCATCTACTTCAACAGAAATCCAGTAGTAAGTTTTCGAAATATCTGCGGTAGAGTTT 1221  
 Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
 QY 1222 GATGGAACTGACCCGAGTAGGAGGAGGATTTACTCTTACCGGGAACGTTGCTTTCCTG 1281  
 Db 281 AspGlyAsnValAlaAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
 QY 1282 AATAATGGAAAAACCTTTGTTCTCAACAATGTTGCTTCTCTCTGTTTACATTTGCTGAAG 1341  
 Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320  
 QY 1342 CAACCAACAGTAGGACAGGCTTCTAATACAGTAATAATTACGAGATGAGGAGGAGCTATC 1401  
 Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
 QY 1402 TTCTGTAAAGATCGTGGCGAA---GCAGGATCCAAATCACTCTGATCAGTTTCTCTTGTAT 1458  
 Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
 QY 1459 GGAGAGGAGTAGTTTTTTTGTAGTCAATGTTAGTCTCTGGAAAGGGGAGCTATTATAT 1518  
 Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380  
 QY 1519 GCCAAAAGCTCTCGGTGCTAACTGTGGCCCTGTACAATTTTTTAAAGAAATATCGCTAAT 1578  
 Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
 QY 1579 GATGGTGGAGCGATTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTTATGGA 1638  
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGlyLeuSerLeuSerAlaAspTyrGly 420

QY 1639 GATATTATTTTCATGGGAACTTTAAAGACAGCAAGAGAAATGCTGCCGATGTTAAT 1698  
DB 421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAsnAlaAspValAsn 440  
QY 1699 GCGCTAACTGTGCTCCACAGCCATTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758  
DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460  
QY 1759 AGAGCTAAAGCAGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818  
DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
QY 1819 AACGAGCCAGCCAGCTCTCCAACTTCTAAATTAACGATGGTGAAGGATACACAGG 1878  
DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyYrThrGly 500  
QY 1879 GATATTGTTTTGCTAAATGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAAGGA 1938  
DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly 520  
QY 1939 AGGATTGTTCTTCGTGAAGCAAAATATCAGTGAATTCCTAAGTCAGACAGGTGG 1998  
DB 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
QY 1999 AGTCTGATATGAAGCTGGGAGTACATGGGATTTGTAACTCCACACCAACCAACAG 2058  
DB 541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
QY 2059 CCTCTGCGCGTAATCAGTTGATCAGCTTTCGAATCTGCAATTTGCTCTTCTCTTTG 2118  
DB 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
QY 2119 TTAGCAAAACAATGAGTACGAATCTCTACAAATCTCCAGCGCAAGATTCATCT 2178  
DB 581 LeuAlaAsnAlaValThrAsnProProThrAsnProAlaGlnAspSerHisPro 600  
QY 2179 GAGTCATTGGTAGCACACTCTGCTGTCTGTTCAATAGTAGGGCTATCTTTTGG 2238  
DB 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePheGlu 620  
QY 2239 GATTGATGATACAGCTATCATAGGTATGATGGTAGGTTCTAATCAAAAATCAAT 2298  
DB 621 AspLeuAspPheThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsp 640  
QY 2299 GTCTGAAATACAGTTAGGACTAAGCCCCAGCTAATGCCCCATCAGATTGACTTA 2358  
DB 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
QY 2359 GGGATGAGATGCTTAAGTATGGCTATCAAGGAGCTGGAAGCTGCGTGGCATCTAAT 2418  
DB 661 GlyAsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
QY 2419 ACAGCAAAATAATGCTCTTATCTCTGAAAGCTACATGCACTAAAGCTGGTATAATCT 2478  
DB 681 ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTyThrLysThrGlyTyAsnPro 700  
QY 2479 GGGCTGACGAGTAGCTCTTTGGTTCCAAATAGTTTATGGGATGCCATTTAGATATA 2538  
DB 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGGATGGGCGCTCTTATTGTCGAGGATTA 2598  
DB 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740  
QY 2599 TGGGTTCTCGAGTTTCGAATTTCTCTATCATGACCGCATGCTTTAGTCCAGGATAT 2658  
DB 741 TrpValSerGlyValSerAsnPhePheTyHisAspArgAspAlaLeuGlnGlyTy 760  
QY 2659 CGGTATATTAGTGGGGTTATTCTCTTAGGAGCAAACTCTCTACTTTGGATCATCATGTTT 2718  
DB 761 ArgTyIleSerGlyTySerLeuGlyAlaAsnSerTyPheGlySerSerMetPhe 780  
QY 2719 GGTCTAGCATTTACCGAAGTATTGTTAGATCTAAAGATTATGTAGTGTGCTGTTCAAT 2778

DB 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAsn 800  
QY 2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCAACAAAGCTTTATGATGGATCTTAT 2838  
DB 801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820  
QY 2839 TTGTTCCGAGATCGGTTTATCCGTTAGCTACCGGTTTGGGAATCAGCATATGAACACC 2898  
DB 821 LeuPheGlyAspAlaPheIleArgAlaSerTyGlyPheGlyAsnGlnHisMetLysThr 840  
QY 2899 TCATATACATTTGCAGAGAGAGGATGTTCTGTTGGGATAATAACTGCTGCTGCGAGAG 2958  
DB 841 SerTyThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
QY 2959 ATTGAGCGGATTAACGATGATTAATCACTCAATCAAGCTCTAATTTGAATGAGTTGGT 3018  
DB 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyLeuAsnGluLeuArg 880  
QY 3019 CTTTTCGTCGCAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAGCGCAT 3078  
DB 881 ProPheValGlnAlaGluPheSerTyAlaAspHisGluSerPheThrGluGluGlyAsp 900  
QY 3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138  
DB 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
QY 3139 TTTGATCAATGTTCTAGTACATCTTAATAATATAGCTTTATGCGGCTTATATCTGT 3198  
DB 921 PheAspArgCysSerSerThrHisProAsnLysTySerPheMetGlyAlaTyIleCys 940  
QY 3199 GATGCTTATCCACCATCTCTGGTACTGAGACAACGCTCTATCCATCCACAGAGACATGG 3258  
DB 941 AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTTTTCATTAGCAAGACATGGAGTTGTGTTAGAGGATCTATGTATGCT 3318  
DB 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyAla 980  
QY 3319 TCTTAAACAAGTAATATAGATATATGCCCATGGAAGATATGATATCGAGATGCTTCT 3378  
DB 981 SerLeuThrSerAsnIleGluValTyGlyHisGlyArgTyGluTyArgAspThrSer 1000  
QY 3379 CGAGCTATGTTTTCAGTGCAGGAGTAGAGTCCGGTTC 3417  
DB 1001 ArgGlyTyGlyLeuSerAlaGlySerLysValArgPhe 1013  
RESULT 7  
ADD43714  
ID ADD43714 standard; protein; 1013 AA.  
XX AC ADD43714;  
XX DT 15-JAN-2004 (first entry)  
XX DE Chlamydia trachomatis immunogenic protein, SEQ ID No 9.  
XX XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
XX KW gene therapy; antibacterial.  
XX OS Chlamydia trachomatis.  
XX PN WO2003049762-A2.  
XX PD 19-JUN-2003.  
XX PF 12-DEC-2002; 2002WO-IB005761.  
XX PR 12-DEC-2001; 2001GB-00029732.  
XX PR 06-AUG-2002; 2002GB-00018233.  
XX PR 14-AUG-2002; 2002GB-00018924.  
XX PA (CHIR-) CHIRON SPA.



QY 1939 AGGATTGTTCTCGTGAAGGCAAAATATCAGTGAATTCCTCAAGTCAGACAGGTGG 1998  
DB |||||  
521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
QY 1999 AGTCGTATATGGAAGCTGGAGTACATGGGATTTTGTAACTCCACACCAACACAG 2058  
DB |||||  
541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
QY 2059 CCTCTGCGCGTAATCAGTTGATACAGCTTTCCTCAATCTGCATTTGCTCTTCTTTG 2118  
DB |||||  
561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
QY 2119 TTAGCAACAATGAGTTAGCAATCTCTACCAATCTCCAGCGCAGATTCATCT 2178  
DB |||||  
581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
QY 2179 CGAGTCATTGGTAGCAACAATCTCGTCTCTGTACAAATAGTAGGCGCTATCTTTTTCAG 2238  
DB |||||  
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGTAGGTCTTAATCAAAAATCAAT 2298  
DB |||||  
621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640  
QY 2299 GTCTCTGAATACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTGACTCTA 2358  
DB |||||  
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
QY 2359 GGGATAGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTGGCGGATCCTAAT 2418  
DB |||||  
661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
QY 2419 ACAGCAAAATATGGTCTTATCTCTGAAGCTACATGAGCTAAACCTGGGTATATCCT 2478  
DB |||||  
681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700  
QY 2479 GGGCTGAGCAGTAGCTCTTGTGTTCCAAATAGTTTATGGGATCCATTTAGATATA 2538  
DB |||||  
701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
QY 2539 CQATCTGCGCATTCAGCAATTCAGCAAGTGGATGGCGCTCTTATTTGTCGAGATTA 2598  
DB |||||  
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
QY 2599 TGGGTTCTGGAGTTTCGAATTTCTTCTATCAGACCGCGATGCTTTAGTCAGGATAT 2658  
DB |||||  
741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
QY 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTACTTTGGATCATCGATGTT 2718  
DB |||||  
761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerMetPhe 780  
QY 2719 GGTCTAGCATTTACCGAAGTATTTGTAGATCTTAAAGATTATGATGTGTCGTTCCCAAT 2778  
DB |||||  
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
QY 2779 CATCATGCTGTCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTCGATCTAT 2838  
DB |||||  
801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
QY 2839 TTGTTCCGGAGATCGGTTTATCCGTCTAGCTAGCGGTTTGGGAATCAGCATATGAACCC 2898  
DB |||||  
821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
QY 2899 TCATATACATTTGCGAGGAGGAGCGATGTTCTGTGGATAATAACTGCTGCTGCGAGAG 2958  
DB |||||  
841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
QY 2959 ATTGAGCGGGATTACCGATTGTGATCTCCATCTAAGCTCTATTGATGAGTTGCGT 3018  
DB |||||  
861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAAGCGAT 3078

DB |||||  
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
QY 3079 CAAGCTCGGCAITTCAGAGCGGCATCTCCCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138  
DB |||||  
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
QY 3139 TTTGATCGATGTTCTAGTACATCCTCTAATAATAATAGCTTTATGGCGCTTATCTGT 3198  
DB |||||  
921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAAGCTCCCTATCCATCAAGAGACATGG 3258  
DB |||||  
941 AspalatyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTTTTCATTTTAGCAAGCATCGAGTGTGGTTAGAGGATCTATGATGCT 3318  
DB |||||  
961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
QY 3319 TCTCTAACAGTAATATAGAAGTATATGGCCATCGAAGATATGATATCGAGATGCTTCT 3378  
DB |||||  
981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
QY 3379 CGAGGCTATGTTTGGTGCAGGAAGTAGAGTCCGGTTC 3417  
DB |||||  
1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013  
RESULT 8  
ADD43786  
ID ADD43786 standard; protein; 1013 AA.  
XX  
AC ADD43786;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 81.  
XX  
KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
KW gene therapy; antibacterial.  
XX  
OS Chlamydia trachomatis.  
XX  
PN WO2003049762-A2.  
XX  
PD 19-JUN-2003.  
XX  
PF 12-DEC-2002; 2002WO-IB005761.  
XX  
PR 12-DEC-2001; 2001GB-00029732.  
PR 06-AUG-2002; 2002GB-00018233.  
PR 14-AUG-2002; 2002GB-00018924.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Grandi G, Ratti G;  
XX  
DR WPI; 2003-532882/50.  
DR N-PSDB; ADD43787.  
XX  
PT New immunogenic composition having a protein or encoding nucleic acid,  
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
PT infection.  
XX  
PS Claim 6; SEQ ID NO 81; 164pp; English.  
XX  
CC The invention relates to a novel immunogenic composition comprising a  
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic  
CC acid comprises any of 131 fully defined amino acid or nucleotide  
CC sequences given in the specification, or has 50% or greater sequence  
CC identity to it, or their fragments. The protein and/or nucleic acid of  
CC the immunogenic composition is useful in the manufacture of a medicament  
CC for the treatment or prevention of infection due to Chlamydia  
CC trachomatis. The infection is treated or prevented by the medicament

CC eliciting an immune response which is specific to a C. trachomatis  
 CC elementary body, or for neutralising C. trachomatis elementary bodies,  
 CC hence the immunogenic composition can be used in creating a vaccine. The  
 CC immunogenic composition can also be used for the diagnosis of C.  
 CC trachomatis infection. The nucleic acids of the immunogenic compositions  
 CC can be used to treat disorders by gene therapy. The immunogenic  
 CC compositions have antibacterial activity. This sequence represents one of  
 CC the 131 C. trachomatis proteins with immunogenic properties of the  
 CC invention.

XX Sequence 1013 AA;

Alignment Scores:  
 Pred. No.: 0 Length: 1013  
 Score: 5131.50 Matches: 983  
 Percent Similarity: 98.8% Conservative: 18  
 Best Local Similarity: 97.0% Mismatches: 11  
 Query Match: 65.1% Indels: 1  
 DB: 7 Gaps: 1

US-10-701-844-1 (1-4435) x ADD43786 (1-1013)

QY 382 ATGCAACCTCTTCCATAGTTCTTTCTTCAATGATCTAGCTTATCTTCTGCTCT 441  
 DB 1 MetGlnThrSerPheHisIysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20  
 QY 442 TTAATGGGGGGATATCGACGAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501  
 DB 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
 QY 502 AGCTTAACGTATCTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
 DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrValPhe 60  
 QY 562 TCTGACGAGAGTTACATTAATAAATCTTGCAATTTCTATGACGCTTTCGCTTTAAGT 621  
 DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
 QY 622 TGTGTTGGAACTTATAGGAGTTTACTGTTTTAGGGAGAGACACTGTTGACTTTC 681  
 DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 QY 682 GAGAACTACGACTTCTCAATAGGGAGCTCTAAGTAATAGCGCTGCTGAGACTG 741  
 DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
 QY 742 TTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTACTTGCCTGA 801  
 DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
 QY 802 CTGCTGCTGCAAGCTAATAAGGGTAGCGACCTCCGACGACCAATCTACCGTCT 861  
 DB 141 LeuProAlaAlaThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
 QY 862 AATGGTACTATTATCTTAAACAGACTTTTCTTACTCAATAGAGAGTTCTCATTC 921  
 DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
 QY 922 TATAGTAAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGAGCTTAAACGTTCAA 981  
 DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
 QY 982 GGAATTAGCAAGCTTTGCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
 DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
 QY 1042 CAAGTAGTACCAGTTTCTGCTATGCTTAAACGAGCTCTTATGCTTTGTCGCGAAT 1101  
 DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
 QY 1102 GTTGACGAGTAAAGGGGGAGGATTTGCTGCTTCCAGATGGGACGAGGAGTGTCA 1161  
 DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlyValSer 260

QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCACAGAAATPACTGCGGTAGAGTTT 1221  
 DB 261 SerSerThrThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
 QY 1222 GATGGAAACGTAGCCCGAGTAGGAGGAGGATTACTCTACGGAGAGGTTGCTTCTCTG 1281  
 DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
 QY 1282 AATAATGCAAAACCTGTTCTCAACAATGTTGCTTCTCTGTTTACATGCTCTCAAG 1341  
 DB 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320  
 QY 1342 CAACCAACAAGTGGACAGGCTTCTAATACAGTAAATAATTACGGAGATGGAGGACTATC 1401  
 DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
 QY 1402 TTCTCTAAGAATGTTGCGCAA---CGAGATCCAATACTCTGGATCAGTTTCTTCTGAT 1458  
 DB 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
 QY 1459 GGAGAGGAGTAGTTTCTTTAGTAGCAATGAGCTGCTGGGAAAGGGGAGCTATTTAT 1518  
 DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380  
 QY 1519 GCCAAAAAGCTCTCGGTTGCTAACTGTGCCCCTGTACAATTTTAAAGGAATATCGCTAAT 1578  
 DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
 QY 1579 GATGTTGAGCCGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638  
 DB 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
 QY 1639 GATATTTATTTTCGATGGGAATCTTAAAGAAACAGCCAAAGAGAGTCTGCCGATGTTAAT 1698  
 DB 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
 QY 1699 GCGCTAACTGTCTCTCACAGCCATTTTCGATGGGATCGGGAGGGGAAATAACGACATTA 1758  
 DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
 QY 1759 AGAGCTAAAGCGGGGATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818  
 DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
 QY 1819 AACCAGCCAGCCGAGTCTTCCAACTTCTAAATAATTAACGATGTCGTAAGGATACACAGG 1878  
 DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyGlyTyrThrGly 500  
 QY 1879 GATATTGTTTTCGCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACGAGGA 1938  
 DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
 QY 1939 AGGATTGTTCTTCGTCGAAAGGCAAAATATATAGTGAATTTCTTAAGTCAGACAGGTTGG 1998  
 DB 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
 QY 1999 AGTCTGTATGAAAGCTGGGATGATGGATTTTGTAACTCCACAAACCCACCAACAG 2058  
 DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
 QY 2059 CCTCTGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCAATTTGTCTCTTCTTCTTCT 2118  
 DB 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580  
 QY 2119 TTACGAAACAATCGAGTTACGAATCTCTTACCAATCTCTTACCAATCTCTTCTTCTTCT 2178  
 DB 581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
 QY 2179 GCGATCTTGGTAGCACACTGCTGTTCTGTTTACAAATTAAGTGGGCTATCTTTTGTAG 2238  
 DB 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProfilePheGlu 620



```
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAATCAAAATCAAT 2298
Db 621 AspLeuAspThrAlaTyArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCCCTGAATACAGTTAGGAGCTAAGCCCGAGCTAATGCCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGATGAGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAGCTGGCGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680
QY 2419 ACAGCAATATATGTCCTTATCTCTGAAAGCTACATCGACTAAACCTGGGTATATCTCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCTGAGCAGTACGTTCTTCTTGGTCCAAATAGTTTATGGGATCCATTTTACATATA 2538
Db 701 GlyProGluA-gValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720
QY 2539 CGATCTGGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAAT 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTCTGGAGTTTCGAAATTTCTTCTATCATGACCGGATGCTTTAGGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTTAGTGGGGTATTTCTTCTAGGACAACTCTCTCTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTTATGTAGTGTGTCGTTCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTGATAGGATCCGTTTATCTATCTATCTACCAAGCTTTATGTCGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTCCGAGATGCTTTATCCGTCGTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTCCAGAGAGAGCGATGTTCTGTCGGATTAATACTGTCGTCGAGAG 2958
Db 841 SerTyrThrPheAlaGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTACCGATTGATGATCTCATCTCATCTCATCTCATCTCATCTCATCTCAT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGATCAATCTTTACAGAGGAAGCGAT 3078
Db 881 PropheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGlyAsp 900
QY 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACATCTCTAATAATATAGCTTTATGCGGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGCACCATCTCTGTCATGACAGAGCTCTCTATCCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGCAGACATGGAGTTGCTGGTTAGAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTTCT 3378
```

```
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGCTATGTTTGGAGTCAGGAAGTAGAGTCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 9
ADD43788
ID ADD43788 standard; protein; 1013 AA.
XX
AC ADD43788;
XX
DT 15-JAN-2004 (first entry)
DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 83.
XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.
XX Chlamydia trachomatis.
OS
PN WO2003049762-A2.
XX
PD 19-JUN-2003.
XX
PF 12-DEC-2002; 2002MO-IB005761.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX
XX (CHIR-) CHIRON SPA.
XX
PI Grandi G, Ratti G;
XX
DR WPI; 2003-532882/50.
XX N-PSDB; ADD43789.
XX
PT New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
XX infection.
PS Claim 6; SEQ ID NO 83; 164pp; English.
XX
CC The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This sequence represents one of
CC the 131 C. trachomatis proteins with immunogenic properties of the
CC invention.
SQ Sequence 1013 AA;
XX
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 7 Gaps: 1
```



```
QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGAGATT 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyA-gSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGAGATTTCGAATTTCTTATCATGACCGCGATGCTTTAGGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPhetYrHisAspA-gAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCTTTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyr-SerLeuGlyAlaAsnSerTyrPheGlySer-SerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTGGTAGATCTAAAGATTATGTAGTGTGCTTCCAAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTACCAACCAAGCTTTATGTGGATCTCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTCCGAGATGCGTTTATCCGTCCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGCGCATGTTGTTGGGATAATAACTGCTCGCTCGAGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTACCGATTCGTGATCTTCCATCTTAAGCTCTATTGAATGAGTTGGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CTTTTCGTCGACGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTTAATAATATAGCTTTATCGCGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACACAGCTCTTATCCCATCAAGAGCATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGCAGACATGAGTGTGGTTAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValAlaArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAATATAGATATATGCGCATATGCGCATATGAGTATCGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGTTGTTAGTGCAGAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 10
ADD43710
ID ADD43710 standard; protein; 1013 AA.
XX AC ADD43710;
XX DT 15-JAN-2004 (first entry)
XX DE Chlamydia trachomatis immunogenic protein, SEQ ID No 5.
XX KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
XX gene therapy; antibacterial.
```

```
OS Chlamydia trachomatis.
XX WO2003049762-A2.
XX PD 19-JUN-2003.
XX PF 12-DEC-2002; 2002WO-IB005761.
XX PR 12-DEC-2001; 2001GB-00029732.
XX PR 06-AUG-2002; 2002GB-00018233.
XX PR 14-AUG-2002; 2002GB-00018924.
XX (CHIR-) CHIRON SPA.
XX Grandi G, Ratti G;
XX WPI; 2003-532882/50.
XX N-PSDB; ADD43711.
XX PT New immunogenic composition having a protein or encoding nucleic acid,
XX useful for diagnosing, preventing and/or treating Chlamydia trachomatis
XX infection.
XX PS Claim 6; SEQ ID NO 5; 164pp; English.
XX CC The invention relates to a novel immunogenic composition comprising a
XX protein or nucleic acid, and an adjuvant, where the protein or nucleic
XX acid comprises any of 131 fully defined amino acid or nucleotide
XX sequences given in the specification, or has 50% or greater sequence
XX identity to it, or their fragments. The protein and/or nucleic acid of
XX the immunogenic composition is useful in the manufacture of a medicament
XX for the treatment or prevention of infection due to Chlamydia
XX trachomatis. The infection is treated or prevented by the medicament
XX eliciting an immune response which is specific to a C. trachomatis
XX elementary body, or for neutralising C. trachomatis elementary bodies,
XX hence the immunogenic composition can be used in creating a vaccine. The
XX immunogenic compositions can also be used for the diagnosis of C.
XX trachomatis infection. The nucleic acids of the immunogenic compositions
XX can be used to treat disorders by gene therapy. The immunogenic
XX compositions have antibacterial activity. This sequence represents one of
XX the 131 C. trachomatis proteins with immunogenic properties of the
XX invention.
XX SQ Sequence 1013 AA;
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 7 Gaps: 1
US-10-701-844-1 (1-4435) x ADD43710 (1-1013)
QY 382 ATGCAACGCTCTTCCATAAGTCTTCTTCAATGATCTAGCTTATTTCTGCTCTT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATATGGGGGATATGACAGCAAAATCATGGTTCTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGTATCATTTCCCTATATCTGTATAGGAGATCCGAGTGGGACTACTGTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGAGGAGAGTTAACTTAAAAAATCTTGACAAATCTTATTCGAGCTTTCGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTGTTGGGAACCTTATTAGGGAGTTTACTGTTTGGGAGGAGGACACTCGTTGACTTTC 681
```

81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy  
682 GAGAACATACGGACTTCTCAAAATGGGCGAGCTTAAGTAATAGCGCTGCTGATGGACTG 741  
Db  
101 GluAenIleArgThrSerThrAenGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120  
Qy  
742 TTACTATTGAGGGTTTAAAGAATTATCTTTTCCAAATTCGAATTCATTACTTGGCGTA 801  
Db  
121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140  
Qy  
802 CTGCTCTGCTCAACGACTAATAAGGATAGCGAGCTCCGACGACAACTATACACGCTCT 861  
Qy  
141 LeuProAlaAlaThrThrAenAenGlySerGlnThrProThrThrSerThrProSer 160  
Db  
862 AATGGTACTATTATTTCTAAACAGATCTTTTGTACTCATATATGAGAAGTCTCATTC 921  
Qy  
161 AenGlyThrIleTySerLysThrAspLeuLeuLeuAenAenGluLysPheSerPhe 180  
Db  
922 TATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981  
Qy  
181 TyrSerAenLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
Db  
982 GGAATTAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTGT 1041  
Qy  
201 GlyIleSerLysLeuCysValPheGlnGluAenThrAlaGlnAlaAspGlyGlyAlaCys 220  
Db  
1042 CAAGTAGTCACAGTTCTCTGCTATGGCTAACGAGGCTCTATTTGCTTTGCTAGCGNAT 1101  
Qy  
221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240  
Db  
1102 GTTCAGAGTAAGAGGGGAGGATCTGCTGTCTAGGATGGCAGCAGGAGTGTC 1161  
Qy  
241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
Db  
1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTTCCAGAAATFACTGCGGTAGAGTTT 1221  
Qy  
261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280  
Db  
1222 GATGGGAAGTACGCCGAGTAGGAGGAGGATTTACTCTCTAGCGGAAGCTTCTCTG 1281  
Qy  
281 AspGlyAenValAlaArgValGlyGlyIleTySerTyGlyAenValAlaPheLeu 300  
Db  
1282 AATAATGAAACCTTTCTCAACAATGTGCTCTCTCTGTTTACATTTGCTGCTAAG 1341  
Qy  
301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyIleAlaAlaGlu 320  
Db  
1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATAATTACGAGATGGAGGAGCTATC 1401  
Qy  
321 GlnProThrAenGlyGlnAlaSerAenThrSerAspAenTyGlyAspGlyGlyAlaIle 340  
Db  
1402 TTCTGTAAAGATGGTGGCAA---GCAGATCCAAATACTCTGGATCAGTTTCTTTGAT 1459  
Qy  
341 PheCysLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAsp 360  
Db  
1459 GGAGGAGGAGTAGTTTCTTTAGTAGCAATGCTAGCTGCTGGAAAGGGGAGCTATTAT 1518  
Qy  
361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTy 380  
Db  
1519 GCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACCAATTTTAAAGAAATATCGCTAAT 1578  
Qy  
381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400  
Db  
1579 GATGCTGAGCGAATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638  
Qy  
401 AspGlyGlyAlaIleTyLeuGlyGlySerGlyGluLeuSerLeuSerAlaAspTyGly 420  
Db  
1639 GATATTATTTTCATGGGAATCTTAAAGAACAGCAAGAACTGCTCCGATGTTAAT 1698  
Qy  
421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440  
Db  
1699 GCGCTAAGTCTGCTCAACGCCATTTTCGATGGGATCGGAGGAGAAATTAACGACATTA 1758  
Qy  
441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyAlaIleThrThrLeu 460  
Db

1759 AGAGCTAAACGAGGGCATCAGATTCTCTTAATGATCCCATCGAGATGGCAACGGAAT 1818  
Qy  
461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAspProIleGluMetAlaAenGlyAen 480  
Db  
1819 AACGAGCAGCGCAGTCTTCCAAACTTCTAAAAATTAACGATGGTGAAGATACACAGGG 1878  
Qy  
481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAspGlyGlyTyThrGly 500  
Db  
1879 GATATTGTTTCTTAATGTAAGCAGTACTTTGTACCAAAATGTTTACGATAGACCAAGGA 1938  
Qy  
501 AspIleValPheAlaAenGlyAenSerThrLeuTyGlnAenValThrIleGlnGlnGly 520  
Db  
1939 AGGATTGTTCTTCTGTAAGGCAAAATTAATCAGTGAATTTCTTAAGTCACACAGGTGGG 1998  
Qy  
521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540  
Db  
1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACAG 2058  
Qy  
541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560  
Db  
2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCAATTTGTCTCTTCTCTTTG 2118  
Qy  
561 ProProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerLeu 580  
Db  
2119 TTACCAACCAATGAGTTCAGATTCCTCCTACCAATCTCCAGCGCAAGATTTCTCATCT 2178  
Qy  
581 LeuAlaAenAenAlaValThrAenProThrAenProProAlaGlnAenSerHisPro 600  
Db  
2179 GCAGTCAATTGGTAGCACAACTGCTGTTCTGTTTCAATTAAGTGGGCTTATCTTTTTCAG 2238  
Qy  
601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
Db  
2239 GATTGGATGATACAGCTTATGATGATGATGCTGGTGGTGGTCTTAATCAAAAATCAAT 2298  
Qy  
621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAenGlnLysIleAsp 640  
Db  
2299 GTCTCTGAAATPACAGTTAGGGACTAAGCCCGCTAATGCCCCATCAGATTTGACTCTA 2358  
Qy  
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660  
Db  
2359 GGGAAATGAGATGCTCAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGATCTTAAT 2418  
Qy  
661 GlyAenGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAen 680  
Db  
2419 ACAGCAATAATGCTTATCTTACTCTGAAAGCTACATGGACTAAACCTGGGTATATCTCT 2478  
Qy  
681 ThrAlaAenAenGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAenPro 700  
Db  
2479 GGGCTGAGCGAGTAGCTTCTTCTGTTCCCAATAGTTTATGGGATCCATTTAGATATA 2538  
Qy  
701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720  
Db  
2539 CGATCTGCGCAATTCAGCAATTAAGCAAGTGTGGATGGGCGCTCTTATTTGTCGAGGATTA 2598  
Qy  
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740  
Db  
2599 TGGGTTTCTGAGGTTTCAATTTCTTATCATGACCGCGATGCTTTTAGGTAGGTAGGATAT 2658  
Qy  
741 TrpValSerGlyValSerAenPheTyHisAspArgAspAlaLeuGlyGlnGlyTy 760  
Db  
2659 CGGTATATTAGTGGGGTTTATCTTCTAGGACCAACTCTCTTTTGGATCATCGATGTTT 2718  
Qy  
761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAenSerTyPheGlySerSerMetPhe 780  
Db  
2719 GGTCTAGCATTTACGAAGTATTTGGTAGATCTAAAGATTTATGATGTGTGTTCCCAAT 2778  
Qy  
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAen 800  
Db  
2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCCAAAGCTTTTATGATGATGATCTAT 2838  
Qy  
801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820  
Db

QY 2839 TTGTCGGAGATCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAAC 2898  
 DB 821 LeuPheGlyAspAlaPheIleAgaIleSerTyGlyPheGlyAenGlnHisMetLysThr 840  
 QY 2899 TCATATACATTTTCAGAGGAGAGCGATGTTCTGGGATATTAATCTGTCTGGCTGGAGAG 2958  
 DB 841 SerTyThrPheAlaGluGluSerAspValArgTIPAspAenAenCysLeuValGlyGlu 860  
 QY 2959 ATTGGAGCGGATTACGGATTGTGATTACTCCATCTAAGCTCTATTGTAATGATGGCT 3018  
 DB 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyLeuAenGluLeuArg 880  
 QY 3019 CCTTCCTGCAAGCTGATTTCTTATGCCGATCATGAATCTTTTACAGAGGAGCGCAT 3078  
 DB 881 ProPheValGlnAlaGluPheSerTyAlaAspHisGluSerPheThrGluGluGlyAsp 900  
 QY 3079 CAAGCTCGGGCATTTCAAGAGCGGACATCTCTAAATCTATCAGTCTCTGTGGAGTAAG 3138  
 DB 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
 QY 3139 TTGATCGATGTTCTAGTACATCCTTAATAATATATAGCTTTATGGCGCTTATCTGT 3198  
 DB 921 PheAspArgCysSerSerThrHisProAenLysTySerPheMetGlyAlaTyIleCys 940  
 QY 3199 GATGCTTATCGACCATCTCTGCTGCTGAGACACGCTCTCTATCCCATCAAGACATGG 3258  
 DB 941 AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTIP 960  
 QY 3259 ACAACAGATGCTTTTCATTTTAGCAGACATGAGTGTGTGTTAGAGATCTATATATGCT 3318  
 DB 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyAla 980  
 QY 3319 TCTCTACAGATATATAGATATATGCGCATCGAGATATGATATGAGATGCTTCT 3378  
 DB 981 SerLeuThrSerAenIleGluValTyGlyHisGlyArgTyGluTyArgAspThrSer 1000  
 QY 3379 CGAGGCTATGTTGTAGTGACGAGGAGTAGATCCGGTTC 3417  
 DB 1001 ArgGlyTyGlyLeuSerAlaGlySerLysValArgPhe 1013  
 RESULT 11  
 ID ADD43706 standard; protein; 1013 AA.  
 XX ADD43706;  
 AC ADD43706;  
 CC 15-JAN-2004 (first entry)  
 DT Chlamydia trachomatis immunogenic protein, SEQ ID No 1.  
 DE immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
 KW gene therapy; antibacterial.  
 KW Chlamydia trachomatis.  
 OS WO2003049762-A2.  
 PN 19-JUN-2003.  
 PD 12-DEC-2002; 2002WO-IB005761.  
 PF 12-DEC-2001; 2001GB-00029732.  
 PR 06-AUG-2002; 2002GB-00018233.  
 PR 14-AUG-2002; 2002GB-00018924.  
 XX (CHIR-) CHIRON SPA.  
 XX Grandi G, Ratti G;  
 PI WPI; 2003-532882/50.  
 XX N-P8DB; ADD43707.  
 DR New immunogenic composition having a protein or encoding nucleic acid,  
 XX

PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis infection.  
 PT Claim 6; SEQ ID NO 1; 164pp; English.  
 XX The invention relates to a novel immunogenic composition comprising a protein or nucleic acid, and an adjuvant, where the protein or nucleic acid comprises any of 131 fully defined amino acid or nucleotide sequences given in the specification, or has 50% or greater sequence identity to it, or their fragments. The protein and/or nucleic acid of the immunogenic composition is useful in the manufacture of a medicament for the treatment or prevention of infection due to Chlamydia trachomatis. The infection is treated or prevented by the medicament eliciting an immune response which is specific to a C. trachomatis elementary body, or for neutralising C. trachomatis elementary bodies, hence the immunogenic composition can be used in creating a vaccine. The immunogenic compositions can also be used for the diagnosis of C. trachomatis infection. The nucleic acids of the immunogenic compositions can be used to treat disorders by gene therapy. The immunogenic compositions have antibacterial activity. This sequence represents one of the 131 C. trachomatis proteins with immunogenic properties of the invention.  
 XX SQ Sequence 1013 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 1013  
 Score: 5131.50 Matches: 983  
 Percent Similarity: 98.8% Conservative: 18  
 Best Local Similarity: 97.0% Mismatches: 11  
 Query Match: 65.1% Indels: 1  
 DB: 7 Gaps: 1  
 US-10-701-844-1 (1-4435) x ADD43706 (1-1013)  
 QY 382 ATGCAACGCTTTCATAGTCTTCTTCATGATCTAGCTATCTTGTGCTCT 441  
 DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTySerCysSer 20  
 QY 442 TTAATCGGGGGATATGACAGAAATCATGTTCTCTCAAGAAATTTACAGTGGGAG 501  
 DB 21 LeuSerGlyGlyGlyTyAlaAlaGluIleMetIleProGlnGlyIleTyAspGlyGlu 40  
 QY 502 ACGTAACTGTATCTTCCCTATCTTATAGAGATCCGAGTGGGACTCTGTTTTT 561  
 DB 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60  
 QY 562 TCTCAGAGAGATTAACTTAAATACTTGACAACTTATTCAGCTTCCTTCTTAAGT 621  
 DB 61 SerAlaGlyGluLeuThrLeuLysAenLeuAspAenSerIleAlaAlaLeuProLeuSer 80  
 QY 622 TGTTTTCGGAATTTAGGGAGTTTCTCTTTTAGGAGAGGACACTCGTTGACTTTC 681  
 DB 81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 QY 682 GAGACATACCGACTTCTACAAATGGGCGACTCTAAGTAATAGCGCTGCTGATGAGCTG 741  
 DB 101 GluAenIleArgThrSerThrAenGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120  
 QY 742 TTTACTATGAGGTTTAAAGATATCTCTTTCATTCGCAATTCATCTACTCCGCTA 801  
 DB 121 PheThrIleGlyPheLysGluLeuSerPheSerAenCysAenSerLeuAlaVal 140  
 QY 802 CTGCTGTCTCAACGACTAATAAGGTAGCAGACTCCGACGACACACTCTACACCGTCT 861  
 DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
 QY 862 AATGCTACTATTATTCTAAACAGATCTTTTGTGTACTCAATAATAGAGAGTTCTCATTC 921  
 DB 161 AenGlyThrIleTySerLysThrAspLeuLeuLeuAenAenGluLysPheSerPhe 180  
 QY 922 TATAGTAATTAGTCTCTGGAGTGGGAGCTATAGATGCTTAAGAGCTTAAAGGTTCAA 981

Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAATTAGCAAGCTTTGTCTCTCCAGAAATATCTGCTCAAGCTCATGGGGAGCTTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACCAAGTTCTCTGCTATGGCTAAAGAGCTCTTATTGCTCTTTAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTCAGGAGTAGAGGGGAGGATTGCTGCTGTTTCAGAGTGGCAGCAGGAGGTGCA	1161
Db	241	ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACGCGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGAACTAGCCGAGTAGGAGGAGGATTACTCTACGGGAACGTGCTTTCCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAAAAACCTTTGTTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu	320
QY	1342	CAACCAAGTCGACAGGCTTCTTAATACAGTAATAATTACGAGATGGAGAGCTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle	340
QY	1402	TTCTGTAAGAATCGTCGCAAA---GCAGGATCAATACTCTCGATCAGTTTCTCTTGAT	1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGGAGTAGTTTCTTTAGTCAATGTAGCTGCTGGAAAGGGGAGCTATTTAT	1518
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr	380
QY	1519	GCCAAAAGCTCTCGTTGCTACTGTGGCCCTGTACAAATTTTAAGGATATCGCTAAT	1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGGTGGAGGATTTATTTAGAGAGATCTGGAGAGCTCAGTTTATCTGCTGATTTGGA	1638
Db	401	AspGlyGlyAlaIleTyrLeuGlyLysSerGlyLysLeuSerLeuSerAlaAspTyrGly	420
QY	1639	GATATTATTTTCGATGGGAATCTTAAAGAACACGCAAGAGAAATGCTGCCGATGTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGCGTAACTGTCTCACAAGCCATTTTCGATGGGATCGGAGGAGGAAATAACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu	460
QY	1759	AGAGCTAAACGAGGCTCAGATTCTCTTAAATGATCCCATCGAGATGGCAACGGAAT	1818
Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACCAAGCCAGCGAGTCTTCCAACTCTTAAATAATTAACGATGCTGAAGGATACACAGG	1878
Db	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly	500
QY	1879	GATATCTTTTGTAAATGGAAGCAGTCTTTGTACCAAAATGTTACGATAGCAGCAAGGA	1938
Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATCTTCTTCGTGAAGGCAAAATATTCAGTGAATCTCTAAGTCACAGAGTGGG	1998
Db	521	ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACCAACCAACAG	2058
Db	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CCTCTCGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCATTTGTCTCTTCTTTG	2118
Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580
QY	2119	TTAGCAAAACAATGCAAGTTACGAATCTCTTACCAATCTCCAGCCGCAAGATCTCATCT	2178
Db	581	LeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	CGAGTCATTGGTAGCACAACTGCTGTTCTGTACAAATTAGTGGGCCCTATCTTTTTCAG	2238
Db	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTTCGATGATACAGCTTATGATAGGTATGATTCGGTAGGTCTTAATCAAAAATCAAT	2298
Db	621	AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTTCGAAATTAACAGTTAGGAGTAAAGCCCCAGCTTAATGCCCCATCAGATTTGACTTA	2358
Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGATCCCTAAGTATGCTATCAAGGAAGCTGGAAGCTGCGTGGGATCCCTAAT	2418
Db	661	GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaIleThrLysThrGlyTyrAsnPro	680
QY	2419	ACAGCAAAATAATGCTCTTATCTCTGAAAGCTACATGAGCTAAAACTCGGTATAATCCT	2478
Db	681	ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro	700
QY	2479	GGCGCTGAGCAGTAGCTTCTTTGTTCCAAATAGTTTATGGGATCCATTTTAGATATA	2538
Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle	720
QY	2539	CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGAGGATTA	2598
Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740
QY	2599	TGGGTTTCTGGAGTTTCGAAATTTCTTCTATCATGACCGCGATGCTTTTAGTCAAGGATAT	2658
Db	741	TyrValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760
QY	2659	CGGTATATTAGTGGGGTTTATCTCTTAGAGCAAACTCTTCTTGGATCATCGATGTTT	2718
Db	761	ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780
QY	2719	GGTCTAGCATTTTACCAAGTATTGCTAGATCTTAAAGATTATGCTAGTGTCTGCTCAAT	2778
Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn	800
QY	2779	CATCATGCTTGATAGGATCCGTTTATCTATCTTACCAACAAGCTTTTATGTGGATCCTAT	2838
Db	801	HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr	820
QY	2839	TTGTTCCGAGATCGGTTTATCCGTGCTAGCTACGGGTTGGGAATCAGCATATGAAAACC	2898
Db	821	LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr	840
QY	2899	TCATATACATTTCCAGAGGAGCGATGTTGTTGGGTAATACTGCTGCTGGCTGGAGAG	2958
Db	841	SerTyrThrPheAlaGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu	860
QY	2959	ATTGGAGCGGATTTACCGATTGTGATTACTTCCATCTAAGCTCTATTATTGATGTTGCGT	3018
Db	861	IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg	880
QY	3019	CCTTTTCGTCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT	3078
Db	881	ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp	900
QY	3079	CAAGCTCGGCATTCAGAGCGGACATCTCTAAATCTATCAGTTCTGCTGGGTGAAG	3138
Db	901	GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys	920





Db 301 AenAenGlyLysThrLeuPheLeuAenValAlaSerProValTyrIleAlaGlu 320  
 QY 1342 CAACCAACAGTGGAGCAGCTCTTAATACGAGTAATAATACGAGATGAGGAGCTATC 1401  
 Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAenAsnTyrGlyAenGlyAlaIle 340  
 QY 1402 TTCTGTAAAGTGGTCGGCAA--GCAGATCAATACTCTGGATCAGTTCCTTTGAT 1458  
 Db 341 PheCysLysAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAen 360  
 QY 1459 GGAGGGAGTAGTTCTTTCTAGTACAAATGTAGCTCTGGGAAAGCGGAGCTATTTAT 1518  
 Db 361 GlyGluGlyValPhePheSerAenValAlaAlaGlyGlyGlyAlaIleTyr 380  
 QY 1519 GCCAAAAGCTCTCGGTGCTTAACCTGGCCCTGTACAATTTTAAAGAAATATCGCTAAT 1578  
 Db 381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400  
 QY 1579 GATGGTGGAGCGATTTATTTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTTATGA 1638  
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420  
 QY 1639 GATATTATTTCCATGGGAATCTTAAAGAAACAGCCAAAGAGAAATGCTGCCGATGTTAAT 1698  
 Db 421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440  
 QY 1699 GCGTAACTGTGTCTCACAGGCCATTTTCATGGGATCGGAGGGGAAATAACGACATTA 1758  
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460  
 QY 1759 AGAGCTAAAGCAGGCGATCAGATCTCTTTAATGATCCCATCCAGATGGCAACGGAAT 1818  
 Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAenProIleGluMetAlaAenGlyAen 480  
 QY 1819 AACAGCCAGCGAGCTCTCAAACTCTTAAATTAACCATGCTGAAGGATACACAGGG 1878  
 Db 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAspGlyGluGlyTyrThrGly 500  
 QY 1879 GATATTGTTTTGCTAATGAAGCAGTACTTTGTACCAAAATGTTACGATAGAGCAAGGA 1938  
 Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyrGlnAenValThrIleGluGlnGly 520  
 QY 1939 AGATGTTCTTCGTGAAAGGCAAAATTAATCAGTGAATCTCTAAGTCAGACAGGTGGG 1998  
 Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540  
 QY 1999 AGTCTGTATATGAAGCTGGAGTACATGGATTTGTTAACTCCACAAACACCAACACAG 2058  
 Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAenPheValThrProGlnProGlnGln 560  
 QY 2059 CCTCTCGCGCTAATCAGTTGATCACGCTTTCCAATCTGCATTTGTCTCTTTCTTTTG 2118  
 Db 561 ProProAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeu 580  
 QY 2119 TTAGCAACAATGCAATGCAATCTCTCCATCCATCCAGCGCAAGATCTCATCCT 2178  
 Db 581 LeuAlaAenAenAlaValThrAenProThrAenProAlaGlnAenSerHisPro 600  
 QY 2179 GCAGTCATTTGTAGCACAACTCGTGTCTGTACAAATAGTGGGCTATCTTTTGTAG 2238  
 Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
 QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATTTGGCTAGTCTTAATCAAAAATCAAT 2298  
 Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspIlePheLeuGlySerAenGlnLysIleAsp 640  
 QY 2299 GTCTGAAATTAAGTATAGGACCTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA 2358  
 Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660  
 QY 2359 GGGATGAGATGCCATGATGCTATCAAGGAGCTGGAAGCTGGTGGGATCCCTAAT 2418

Db 661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAen 680  
 QY 2419 ACAGCAAAATATGTCCTTATCTGAAAGCTACATGACTATAAAGCTGGGTATATCTCT 2478  
 Db 681 ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTyrPheThrGlyTyrAenPro 700  
 QY 2479 GGGCTCGAGCAGTAGTCTTTTGGTTCCTCAATAGTCTTATGGGATCCATTTAGATATA 2538  
 Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTyrGlySerIleLeuAspIle 720  
 QY 2539 CGATCTGGCATTCAGCAATTCACCAAGTGGATGGCGCTCTTATGTTCGAGGATTA 2598  
 Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
 QY 2599 TGGGTTCTGGAGTTTCGAATTTCTTATCATGACCGGATGCTTTAGCTCAGGATAT 2658  
 Db 741 TrpValSerGlyValSerAenPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
 QY 2659 CGGTATATTAGTGGGGTTTATCTTTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718  
 Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780  
 QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTCTGTTCCAAT 2778  
 Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800  
 QY 2779 CATCATGCTGCATAGGATCCGTTTATCTTATCTACCCAAACAGCTTTATGTGATCCTAT 2838  
 Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
 QY 2839 TTGTTCCGAGATCGCTTTATCCGTGCTAGCTACGGTTTGGGAATCAGCATATGAAAAAC 2898  
 Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840  
 QY 2899 TCATATACATTTTCAGAGAGGAGCGATGTTGTTGGGATAATACTGTCTGGCTGAGAG 2958  
 Db 841 SerTyrThrPheAlaGluSerAspValArgTyrAspAenAenCysLeuValGlyGlu 860  
 QY 2959 ATTGGAGCGGATACCGATGTTGATTCTCCATCTAAGCTCTTATTTGAATGATGTTGCGT 3018  
 Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880  
 QY 3019 CCTTTCTGTCAGCTCAGTCTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT 3078  
 Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
 QY 3079 CAAGCTCGGCATTCAGAGCGGACATCTCTTAATCTATCATCTCTGTTCTGAGTCAAG 3138  
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920  
 QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGCGGCTTATATCTGT 3198  
 Db 921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940  
 QY 3199 GATGCTTATCGCACCATCTCTGCTACTGAGCAACGCTCTCTATCCCATCAAGAGACATGG 3258  
 Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTyr 960  
 QY 3259 ACAACAGATGCCCTTTCATTTAGCAAGACATGAGTGTGTGTGTAGAGATCTATGTATGCT 3318  
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
 QY 3319 TCTCTAACAGTAAATAGAGTATATGGCCATGAGGATATGATGATCGAGATGCTTCT 3378  
 Db 981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
 QY 3379 CGAGGCTATCGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417  
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013  
 RESULT 13  
 AEA19078  
 ID AEA19078 standard; protein; 1013 AA.

XX AEA19078;  
 XX AC  
 XX 28-JUL-2005 (first entry)  
 XX DB  
 XX Chlamydia trachomatis protein - SEQ ID 81.  
 XX DE  
 XX KW chlamydia trachomatis infection; antibacterial; vaccine.  
 XX OS  
 XX Chlamydia trachomatis.  
 XX PN  
 XX US2005106162-A1.  
 XX PD  
 XX 19-MAY-2005.  
 XX PF  
 XX 16-DEC-2004; 2004US-00498327.  
 XX PR  
 XX 12-DEC-2001; 2001GB-00029732.  
 XX PR  
 XX 06-AUG-2002; 2002EP-00182233.  
 XX PR  
 XX 14-AUG-2002; 2001EP-00218924.  
 XX PR  
 XX 12-DEC-2002; 2002WO-IB005761.  
 XX PA  
 XX (GRAN/) GRANDI G.  
 XX PA  
 XX (RATT/) RATTI G.  
 XX PI  
 XX Grandi G, Ratti G;  
 XX DR  
 XX WPI; 2005-354777/36.  
 XX DR  
 XX N-PSDB; AEA19079.  
 XX XX  
 XX New immunogenic composition comprising Chlamydia trachomatis protein or  
 XX PT its encoding nucleic acid, useful for immunizing against, or treating or  
 XX PT preventing chlamydial infection.  
 XX PS  
 XX Claim 13; SEQ ID NO 81; 90pp; English.  
 XX CC  
 XX The invention comprises an immunogenic composition for the treatment or  
 XX CC prevention of a Chlamydia trachomatis infection. The immunogenic  
 XX CC composition contains a Chlamydia trachomatis protein or its encoding  
 XX CC nucleic acid. The immunogenic composition of the invention is useful for  
 XX CC treating or preventing chlamydial infection. The present amino acid  
 XX CC sequence represents a Chlamydia trachomatis protein of the invention.  
 XX CC NOTE: The present sequence is not shown in the specification, but was  
 XX CC obtained from the USPTO website -  
 XX CC seqdata.uspto.gov/sequence.html?DocID=20050106162.  
 XX SQ  
 XX Sequence 1013 AA;

Alignment Scores:  
 Pred. No.: 0 Length: 1013  
 Score: 5131.50 Matches: 983  
 Percent Similarity: 98.8% Conservative: 18  
 Best Local Similarity: 97.0% Mismatches: 11  
 Query Match: 65.1% Indels: 1  
 DB: 9 Gaps: 1

US-10-701-844-1 (1-4435) x AEA19078 (1-1013)

QY 382 ATGCAACGCTCTTCCATAGTTCTTTCTTCAATGATCTAGCTTATTCTGCTGCT 441  
 DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20

QY 442 TTAATGGGGGGATATCGACAGAAATCATGGTTCCTCAAGGAATTACATGGGGAG 501  
 DB 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40

QY 502 AGCTTAACGTATCTTCCCTATCTGTATAGGAGATCCGAGTGGGACTAGTGT 561  
 DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60

QY 562 TCTGCAGGAGAGTTAACTTAAAAAATCTTGCAATTTCTATGACGCTTTCCTTAAGT 621  
 DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80

QY 622 TGTTTGGGAACCTATTAGGAGTTTACTGTTTGGGAGGACACACTCGTTGACTTTC 681  
 DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100

QY 682 GAGAACATACGACTCTTACAAATGGGCGACTCTTAAGTAATAGCGCTGCTGATGACTG 741  
 DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120

QY 742 TTTACTATTGAGGGTTTAAAGAAATTATCTTTTCCAAATTCGAATTCATTCTTCCGTA 801  
 DB 121 PheThrIleGluGlyPheLysLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140

QY 802 CTGCTGCTGCAACGACTAATAAGGTTAGCAGACTCCGACGACAACTACACCGTCT 861  
 DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160

QY 862 AATGTTACTATTATTCTTAAACAGACTCTTTTGTGTACTCAATATAGAAAGTTCTCATTC 921  
 DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180

QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA 981  
 DB 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200

QY 982 GGAATTAGCAAGCTTTGTCTCTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
 DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220

QY 1042 CAAGTAGTCACAGTTTCTCTGCTATGCTTAAACAGAGGCTCTTATGCTTGTAGCGAAT 1101  
 DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240

QY 1102 GTTCAGAGTAAGAGGGGAGGATTGCTGCTGTTCAAGATGGGCACAGAGGAGTGCA 1161  
 DB 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260

QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACCTGCGGTAGAGTTT 1221  
 DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280

QY 1222 GATGGAAACGTAGCCCGAGTAGGAGGAGGATTACTCTACGGGAAAGTTGCTTTCTCTG 1281  
 DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300

QY 1282 AATAATGGAAAAACCTTGTCTCAACAAATGTTGCTTCTCTGTTTCAATTCATGCTGCTAAG 1341  
 DB 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320

QY 1342 CAACCAACAGTGGACAGGCTTCTTAATACGAGTAATAATACGGAGATGGGAGGACTATC 1401  
 DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340

QY 1402 TTCTGTAAAGATGGTGCGCAA---GCAGGATCCAAATACTCTGATCAGTTTCTTTGAT 1458  
 DB 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360

QY 1459 GGAGAGGAGTAGTGTCTTTTAGTAGCAATGTAGTCTGTGGGAAAGGGGAGCTATTTAT 1518  
 DB 361 GlyGluGlyValAlaPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380

QY 1519 GCCAAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAATTTTAAAGGAATATCGCTAAT 1578  
 DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400

QY 1579 GATGCTGAGCGCAATTTATTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTATGA 1638  
 DB 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420

QY 1639 GATATATTTTCGATGGGAATCTTAAAGAACAGCAAGAGAAATGCTGCCCATGTTAAT 1698  
 DB 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440



DR MPI; 2005-354777/36.  
 DR N-PSDB; ABA19003.

XX New immunogenic composition comprising Chlamydia trachomatis protein or  
 PT its encoding nucleic acid, useful for immunizing against, or treating or  
 PT preventing chlamydial infection.

XX Claim 13; SEQ ID NO 5; 90pp; English.

XX The invention comprises an immunogenic composition for the treatment or  
 CC prevention of a Chlamydia trachomatis infection. The immunogenic  
 CC composition contains a Chlamydia trachomatis protein or its encoding  
 CC nucleic acid. The immunogenic composition of the invention is useful for  
 CC treating or preventing chlamydial infection. The present amino acid  
 CC sequence represents a Chlamydia trachomatis protein of the invention.  
 CC NOTE: The present sequence is not shown in the specification, but was  
 CC obtained from the USPTO website -  
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.

XX Sequence 1013 AA;

Alignment Scores:

Pred. No.:	0	Length:	1013
Score:	5131.50	Matches:	983
Percent Similarity:	98.8%	Conservative:	18
Best Local Similarity:	97.0%	Mismatches:	11
Query Match:	65.1%	Indels:	1
DB:	9	Gaps:	1

US-10-701-844-1 (1-4435) x ABA19002 (1-1013)

QY	382	ATGCAAACTCTTTCATAGTTCTTCTTCAATGATCTAGCTATCTTCTGCTCT	441
Db	1	MetGlnThrSerPheHieIysPheLeuSerMetIleLeuAlaIyrSerCysSer	20
QY	442	TTAATATGGGGGATATGCACAGAAATCATGGTTCTCTCAAGGAATTTACGATGGGAG	501
Db	21	LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu	40
QY	502	ACGTTAACTGTATCTATCTCCCTATCTGTTATAGGAGATCCGAGTGGGACTCTGTTTT	561
Db	41	ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTGAGGAGAGTTAACAATAAAATCTTGACAAATCTATTTGCGAGCTTGGCTTTAAGT	621
Db	61	SerAlaGlyGluLeuThrLeuIysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
QY	622	TGTTTTGGGAATTTATAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC	681
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
QY	682	GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTG	741
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
QY	742	TTTACTATTGAGGTTTTAAAGAAATATCTCTTTTCCAATGTCATTTACTTGGCGTA	801
Db	121	PheThrIleGluGlyPheIysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCGCTGCTGCAAGCACTAATAAGGCTAGCCAGACTCCGACGACACATCTACACGCTCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	AATGGTACTATTATTCTAAAAACAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTC	921
Db	161	AsnGlyThrIleTyrSerIysThrAspLeuLeuLeuLeuAsnAsnGluIysPheSerPhe	180
QY	922	TATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTTAAGGCTTAACGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaIysSerLeuThrValGln	200
QY	982	GGAAATTTAGCAAGCTTTGCTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT	1041

Db	201	GlyIleSerIysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACAGTTTCTCTGCTATGCTTAACGAGGCTCTATTTGCTTTGTAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTCCAGGAGTAACAGAGGGGAGGATTGCTGCTTTCAGGATGGCGACAGGAGTGTC	1161
Db	241	ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTCTCAACAGAGATCCAGTAGTAGTTTTCAGAAATATCTGCGGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGAACTAGTCCCGAGTAGGAGGAGGATTACTCTCTACGGGAAGTTGCTTCTCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAAAACTGTTTCTCAACAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT	1341
Db	301	AsnAsnGlyIysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu	320
QY	1342	CAACCAACAGTGGACAGGCTTCTAATACGAGTAAATAATACGAGATGGAGGAGCTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle	340
QY	1402	TTCTGTAAAGATGTTGGCCAA---GCAGGATCCAAATACTCTGATCAGTTTCTCTTCAAT	1458
Db	341	PheCysIysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGAGTAGTTTCTTTAGTAGCAATCTAGCTGCTGGGAAAGGGGAGCTATTTAT	1518
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIysGlyAlaIleTyr	380
QY	1519	GCCAAAAAGCTCTCGTTTCTGCTAACTGTGGCCCTGTACAAATTTTAAAGAAATATCGCT	1578
Db	381	AlaIysIysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGTGAGCGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATATATGA	1638
Db	401	AspGlyGlyAlaIleTyrLeuGlyIleSerGlyGluLeuSerLeuSerAlaAspTyrGly	420
QY	1639	GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCGCATGTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuIysArgThrAlaIysGluAsnAlaAlaAspValAsn	440
QY	1699	GGCGTAACTGTCTCTCAAGCCATTTTCGATGGGATCGGAGGGGAAATAACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIysIleThrThrLeu	460
QY	1759	AGAGCTAAAGCAGGGCATCAGATCTCTTTAATGATCCATCCAGATCGCAACCGAAT	1818
Db	461	ArgAlaIysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACGAGCAGCGCAGCTTCTCCAACTTCTTAAATAATTAACGATGGTGAAGGATACACAGG	1878
Db	481	AsnGlnProAlaGlnSerSerGluProLeuIysIleAsnAspGlyGlyGlyTyrThrGly	500
QY	1879	GATATTGTTTTGCTTAATGAAGCAGTACTTTGTACCAAAATGTTTACGATAGACAGA	1938
Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATTGTTCTTCTGTAAGCAAAATATCAGTGAATTTCTTAAGTCAGACAGGTGG	1998
Db	521	ArgIleValLeuArgGluIysAlaIysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGTATATGAAGCTGGAGTACATGGGATTTTGTAACTCCCAACACCAACACAG	2058
Db	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CCTCTGCGCGCTTAATCAGTTGATCAGCTTTTCCAAATCTGCAATTTGCTCTTTCTTTG	2118
Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu	580



DB:	9	Gaps:	1
US-10-701-844-1 (1-4435) x ABA19080 (1-1013)			
QY	382	ATGCAGAGCTCTTCCATAAGTCTCTTCTCAATGATTCTAGCTTATTCTTGCTGCTCT	441
Db	1	MetGlnThrSerPheHisLeuPhePheLeuSerMetLeuLeuAlaIleTySerCysSer	20
QY	442	TTAAATGGGGGGATATGCAGAGAAATCATGTTCTCAAGGAATTTACGATGGGAG	501
Db	21	LeuSerGlyGlyTyAlaAlaGluIleMetIleProGlnGlyIleTyAspGlyGlu	40
QY	502	ACGTTAACTGATCATTTCCCTATCTCTTATAGGAGATCCGATGGGACTACTGTTTT	561
Db	41	ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTCGAGAGAGTTAAACATTAATAATCTTCACAAATCTATTGACGCTTCCTTTAAGT	621
Db	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
QY	622	TGTTTGGGAATTTATTAGGAGTTTACTGTGTTTGGGAGAGGACACTCGTTGACTTTC	681
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyAArgGlyHisSerLeuThrPhe	100
QY	692	GAGAACTACGGAATCTTCAAAATGGGCGAGCTCTAAGTAATAGCGCTGTGATGGACTG	741
Db	101	GluAsnIleAArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
QY	742	TTTACTATTGAGGTTTAAAGAAATATCTCTTCCAAATTCGAATTCATTTACTTCCGCTA	801
Db	121	PheThrIleGluGlyPheLeuGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGTCTGCAACCACTAATAAGGTAGCAGACTCCGACGACCAACATCTACACCGTCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	AATGTAATATTATTCTTAAACAGATCTTTTGTACTCAATTAATAGAGAGTTCTCATTC	921
Db	161	AsnGlyThrIleTySerLysThrAspLeuLeuLeuLeuLeuAsnAsnGluLysPheSerPhe	180
QY	922	TATAGTAATTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAAATAGCAAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACCAAGTTCTCTGCTATGGCTAACGAGGCTCTATTGGCTTTGTAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTCGAGGATGAAGGGGGAGTTCGTCTGCTTCCAGATGGCGCAGGGAGTGCA	1161
Db	241	ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTCGCGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGGAACGTACCGGAGTAGGAGGAGATTACTCTCAAGGAAAGCTGCTTTTCCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAAAACTTTCTCAACAATGTGCTTCTCCCTGTTTACATTCGCTGCTAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu	320
QY	1342	CAACCAACAGGTGGACAGGCTTCTAATACAGTAATAATACGAGATGCGAGGACTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle	340
QY	1402	TTCTGTGAAGATGGTGGCAA---GCAGATCCAACTCTGGATCAGTTCTCTTTGAT	1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGGAGGAGTAGTTTTCTTTTAGTACAAATGTAGCTCTGGGAAAGGGAGCTATTAT	1518
Db	361	GlyGluGlyValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTy	380
QY	1519	GCCAAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGTAAT	1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATCGTCGAGCGATTTATTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA	1638
Db	401	AspGlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyGly	420
QY	1639	GATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAGAAATGTCGCGATGTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGCTAACTGTCTCTCACAGCCATTTTCGATGGGATCGGAGGAGGAAATTAACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu	460
QY	1759	AGACTAAAGCAGCGCATCATGATCTCTTAAATGATCCCATCGAGATCGCAACCGAAAT	1818
Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACCAAGCAGCGCATCTTCCAAATCTTAAAAATTAACGATGGTGAAGGATACACAGGG	1878
Db	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyThrGly	500
QY	1879	GATATTGTTTGTCTAATGGAAGCAGTACTTTGTGACCAAAATGTTACGATAGCAAGGA	1938
Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyThrGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATGCTCTCTCTGAAAAGGCAAAATATCAGTGAATCTCTAAGTCACAGAGTGGG	1998
Db	521	ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACAG	2058
Db	541	SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CCTCTGCGCTAATCAGTTGATCAGCTTCCAACTGCAATTTGCTCTCTTCTCTCTTG	2118
Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580
QY	2119	TTAGCAACAATGCAGTTACCAATCTCTACCAATCTCTCCAGGCGCAAGATTTCTCATCT	2178
Db	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	GCAGTCAATGTGTAGCACAACTGCTGTTCTGTACAATTAGTGGGCTCTATCTTTTTCAG	2238
Db	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTTGGATGATACAGCTTATGATAGTATGATTTGGCTAGGTTCTAATCAAAAATCAAT	2298
Db	621	AspLeuAspAspThrAlaTyAspArgTyAspTyLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTCTCGAAATTTACAGTTAGGAGCTAAGCCCGCTAATGCCCATCAGATTTGACTCTA	2358
Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGATGCTCAAGTATGGCTTATCAAGGAAGCTGGAAGCTTGGCGGATCTTAAT	2418
Db	661	GlyAsnGluMetProLysTyGlyTyGlnGlySerTyLysLeuAlaTyLysAspProAsn	680
QY	2419	ACAGCAAAATATGTCCTTATCTCTGAAAGCTACATGCAATAAACTGGGTATTAATCTCT	2478
Db	681	ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTyLysThrTyLysThrGlyTyAsnPro	700
QY	2479	GGGCTGAGCGAGTAGTCTCTTTGGTCCCAATAGTTTATGGGATCCATTTTAGATATA	2538

Db 701 GlyProGluArgValAlaSerLeuValProHsenSerLeuTyrGlySerIleLeuAspIle 720  
QY 2539 CGATCTGGCAATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTGAGGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
QY 2599 TGGGTTTCTGGAGTTTCCGAATTTCTTATCATGACCGCGATGCTTTAGGTCAGGATAT 2658  
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspA-GAspAlaLeuGlyGlnGlyTyr 760  
QY 2659 CGGTATATTAGTGGGGTATTCTCTTAGGAGCAAACTCTCTACTTTGGATCATCGATGTTT 2718  
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
QY 2719 GGTCTAGCATTTACGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGTCCTCAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
QY 2779 CATCATGCTTGATAGGATCCGTTTATCTATCCCAAGCTTTATGTGATGCTCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
QY 2839 TTGTTCCGAGATCGCTTTATCGTCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898  
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
QY 2899 TCATATACATTTGCGAGAGAGCGATGTTGCTGGGATTAATACTGTCTGCTGCGAGAG 2958  
Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
QY 2959 ATTGAGCGGATTCACGATTCGATCTCTCATCTTAAGCTCTTATTTGAATGAGTTCGT 3018  
Db 861 IleGlyValGlyLeuProIleThrProSerLysLeuTyrLeuAsnGlnLeuArg 880  
QY 3019 CCTTTTCGTCAGAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluArg 900  
QY 3079 CAAGCTCGGCATTCAGAGCGGACATCTCCTTAATCTATCATCTGCTGCTGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
QY 3139 TTTGATCGATGTTCTAGTACATCTCTATTAATATAGCTTTATGCGCGCTTATCTCT 3198  
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
QY 3199 GATGCTTATCGACCATCTCTGCTACTGAGACAACTCTCTATCCCATCAAGAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATCGCTTTCATTAGCAAGACATGGAGTTGTGGTAGAGATCTATGTATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
QY 3319 TCTCTAACAGTATATAGAGTATATGCCATGGAAGTATGAGTATCAGTATCAGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
QY 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013  
RESULT 16  
ID AEA19006  
X AC AEA19006 standard; protein; 1013 AA.  
X AC AEA19006;  
X XT 28-JUL-2005 (first entry)  
X DE Chlamydia trachomatis protein - SEQ ID 9.  
X XX chlamydia trachomatis infection; antibacterial; vaccine.  
KW

XX Chlamydia trachomatis.  
OS US2005106162-A1.  
PN 19-MAY-2005.  
PD 16-DEC-2004; 2004US-00498327.  
PF 12-DEC-2001; 2001GB-00029732.  
PR 06-AUG-2002; 2002EP-00182233.  
PR 14-AUG-2002; 2001EP-00218924.  
PR 12-DEC-2002; 2002WO-IB005761.  
XX (GRAN/) GRANDI G.  
PA (RATT/) RATTI G.  
PI Grandi G, Ratti G;  
XX WPI: 2005-354777/36.  
XX N-PSDB; AEA19007.  
PT New immunogenic composition comprising Chlamydia trachomatis protein or  
its encoding nucleic acid, useful for immunizing against, or treating or  
preventing chlamydial infection.  
PS Claim 13; SEQ ID NO 9; 90pp; English.  
XX The invention comprises an immunogenic composition for the treatment or  
prevention of a Chlamydia trachomatis infection. The immunogenic  
composition contains a Chlamydia trachomatis protein or its encoding  
nucleic acid. The immunogenic composition of the invention is useful for  
treating or preventing chlamydial infection. The present amino acid  
sequence represents a Chlamydia trachomatis protein of the invention.  
CC NOTE: The present sequence is not shown in the specification, but was  
obtained from the USPTO website -  
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.  
XX SQ Sequence 1013 AA;  
Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5123.50 Matches: 979  
Percent Similarity: 98.8% Conservative: 22  
Best Local Similarity: 96.6% Mismatches: 11  
Query Match: 65.0% Indels: 1  
DB: 9 Gaps: 1  
US-10-701-844-1 (1-4435) x AEA19006 (1-1013)  
QY 382 ATGCAACGCTCTTCCATTAAGTCTCTTCAATGATTCTAGCTTATTCTGCTCTCT 441  
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20  
QY 442 TTAAATGGGGGGGATATGACAGCAAAATCATGGTTCTCTCAAGGAATTTACGATGGGAG 501  
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
QY 502 ACCTTAACGTATCATTTCCCTATATCTGTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGCGAGGAGTTAACTTAATAAATCTTGCAATTCATTCTTATTCACGCTTTGCTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaAlaLeuProLeuSer 80  
QY 622 TGTGTTGGGAACCTTATTAGGAGTTTACTGTGTTTGGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGAACATACGACTTCTTCAAAATGGGCGAGCTCTAAGTAATACGCTGCTGATGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120





Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnCysLeuValGlyGlu 860  
 Qy 2959 ATTGAGCGGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGAATGAGTTGGT 3018  
 Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
 Qy 3019 CCTTCGTGCAAGCTGAGTTTCTTATCCCATCATGAATCTTTTACAGAGCAAGCGCAT 3078  
 Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
 Qy 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTTGGAGTGAAG 3138  
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
 Qy 3139 TTTGATCATGTTCTAGTACATCATCTTAATAATATAGCTTTATGCGCGCTATATCTGT 3198  
 Db 921 PheAspArgCysSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
 Qy 3199 GATGCTTATCGCACCATCTCTGGTACTCAGACAACGCTCTATCCCATCAAGACATGG 3258  
 Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
 Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGTATGCT 3318  
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
 Qy 3319 TCTTAAACAAGTAATATAGAGTATATGCGCCATGGAAGATATGATGATCGAGATCCTTCT 3378  
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
 Qy 3379 CGAGGCTAGTTTTCAGTGCAGGAGTAGAGTCCCGTTC 3417  
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 17  
 ID AEA19090  
 AC AEA19090 standard; protein; 1013 AA.  
 AC AEA19090;  
 DT 28-JUL-2005 (first entry)  
 DE Chlamydia trachomatis protein - SEQ ID 93.  
 KW chlamydia trachomatis infection; antibacterial; vaccine.  
 OS Chlamydia trachomatis.  
 FN US2005106162-A1.  
 PD 19-MAY-2005.  
 PF 16-DEC-2004; 2004US-00498327.  
 PR 12-DEC-2001; 2001GB-00029732.  
 PR 06-AUG-2002; 2002EP-00182233.  
 PR 14-AUG-2002; 2001EP-00218924.  
 PR 12-DEC-2002; 2002WO-IB005761.  
 PA (GRAN/) GRANDI G.  
 PA (RATT/) RATTI G.  
 XX Grandi G, Ratti G;  
 XX WPI: 2005-354777/36.  
 DR N-PSDB; AEA19091.  
 XX New immunogenic composition comprising Chlamydia trachomatis protein or  
 PT its encoding nucleic acid, useful for immunizing against, or treating or  
 PT preventing chlamydial infection.  
 XX Claim 13; SEQ ID NO 93; 90pp; English.

XX The invention comprises an immunogenic composition for the treatment or  
 CC prevention of a Chlamydia trachomatis infection. The immunogenic  
 CC composition contains a Chlamydia trachomatis protein or its encoding  
 CC nucleic acid. The immunogenic composition of the invention is useful  
 CC treating or preventing chlamydial infection. The present amino acid  
 CC sequence represents a Chlamydia trachomatis protein of the invention.  
 CC NOTE: The present sequence is not shown in the specification, but was  
 CC obtained from the USPTO website -  
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.  
 XX  
 SQ Sequence 1013 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 1013  
 Score: 5123.50 Matches: 979  
 Percent Similarity: 98.8% Conservative: 22  
 Best Local Similarity: 96.6% Mismatches: 11  
 Query Match: 65.0% Indels: 1  
 DB: 9 Gaps: 1  
 US-10-701-844-1 (1-4435) x AEA19090 (1-1013)  
 QY 382 ATGCAACGCTCTTCCATAGTCTCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT 441  
 Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysCysSer 20  
 QY 442 TTAATGGGGGGGATATGACGAGAAATCATGTTCTCAAGGAATTTACGATGGGAG 501  
 Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
 QY 502 ACGTTAACTGATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
 QY 562 TCTCAGGAGAGTTAACTTAAATACTTGACAAATCTTGACAAATCTTGACGCTTTTAACT 621  
 Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaAlaLeuProLeuSer 80  
 QY 622 TGTITTTGGAACTTATTAGGGAGTTTACTGTTTGGGAGGAGGACACTCGTTGACTTTC 681  
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 QY 682 GAGAACATACGACCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTCATGACTG 741  
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
 QY 742 TTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTTGCAATTCATTACTTCCGCTA 801  
 Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
 QY 802 CTGCTGCTGCACGACTAATAAGGGTAGCCAGACTCCGAGCAGACACTTACACCGCTCT 861  
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
 QY 862 AATGGTACTATTATTCTTAAACAGATCTTTTGTACTCAATAATAGAGAAATTTCTCATTC 921  
 Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGlnLysPheSerPhe 180  
 QY 922 TATAGTAATTTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981  
 Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
 QY 982 GGAATTAGCAGCTTTGTGCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTGCT 1041  
 Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
 QY 1042 CAAGTAGTCACAGTTTCTCTGCTATGGCTAACAGAGGCTCTTATGCTTTGTAGCGAAT 1101  
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
 QY 1102 GTTCAGGAGTAAAGAGGGGAGGATGCTGCTGTTTCAAGATGGGAGGAGTGTCA 1161

Db 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAaspGlyGlnGlnGlyValSer 260  
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTCGGGTAGAGTTT 1221  
Db 261 SerSerThrThrGluAaspProValValSerPheSerArgAenThrAlaValGluPhe 280  
Qy 1222 GATGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTTCCTG 1281  
Db 281 AspGlyAasnValAlaAlaGValGlyGlyIleTyrSerTyrGlyAasnValAlaPheLeu 300  
Qy 1282 AATAATCGAAAAACCTTGTTCCTCAACAATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1341  
Db 301 AasnAasnGlyAasnValAlaAasnValAlaAasnValAlaAasnValAlaAasnValAla 320  
Qy 1342 CAACCAACAGTGCAGGCTTCTAATACGAGTAATTAATACGAGATGAGGAGCTATC 1401  
Db 321 GlnProThrAasnGlyGlnAlaSerAasnThrSerAasnTyrGlyAaspGlyAlaIle 340  
Qy 1402 TTCTGTAAAGATGTCGCA--GCAGGATCCCAATTAATCTGGATCAGTTTCTCTTGTAT 1458  
Db 341 PheCysAasnGlyAlaGlnAlaAlaGlySerAasnAasnSerGlySerValSerPheAasp 360  
Qy 1459 GGAGAGGAGTAGTTTCTTTTAGTAGCAATGATGCTCTGGAAAGGGGAGCTATTAT 1518  
Db 361 GlyGluGlyValValPhePheSerSerAasnValAlaAlaGlyAasnGlyAlaIleTyr 380  
Qy 1519 GCCAAAAGCTCTCGTGTCTAATCTGGCCCTGTACATTTTAAAGGAATATCGCTAAT 1578  
Db 381 AlaIysIysLeuSerValAlaAasnCysGlyProValGlnPheLeuGlyAasnIleAlaAasn 400  
Qy 1579 GATGGTGGAGGATTTATTTAGGAGATCTCGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
Db 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlnLeuSerLeuSerAlaAaspTyrGly 420  
Qy 1639 CATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAAATGTCGCGATGTTAAT 1698  
Db 421 AspIleIlePheAaspGlyAasnLeuIysArgThrAlaIysGluAasnAlaAaspValAasn 440  
Qy 1699 GCGCTAATGTCCTCAACAGCCATTTCGATGGATCGGAGGGGAAATTAACGACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460  
Qy 1759 AGAGCTAAGCAGGCGATCAGATTCCTTTAATGATCCATCGAGATGCGCAACCGAAAT 1818  
Db 461 ArgAlaIysAlaGlyHisGlnIleLeuPheAasnAaspProIleGluMetAlaAasnGlyAasn 480  
Qy 1819 AACCAAGCCAGCGAGCTCTTCCAAACTCTTAAATAATTAACGATGGTGAAGGATACACGGG 1878  
Db 481 AasnGlnProAlaGlnSerSerGluProLeuIysIleAasnAaspGlyGluGlyTyrThrGly 500  
Qy 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACGAAGA 1938  
Db 501 AspIleValPheAlaAasnGlyAasnSerThrLeuTyrGlnAasnValThrIleGlnGlnGly 520  
Qy 1939 AGGATTGTTCTCGTGAAGGCAAAATTTATCAGTGAATTCCTAAGTCAGACAGCTGGG 1998  
Db 521 ArgIleValLeuArgGlnIysAlaIysLeuSerValAasnSerLeuSerGlnThrGlyGly 540  
Qy 1999 AGTCGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTTCCCAACCAACCAACAG 2058  
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAaspPheValThrProGlnProGlnGln 560  
Qy 2059 CTTCTGCGCGTAATCAGTTGATCAGCTTTCGATCTCGATTTGCTCTTCTCTCTCTCTCT 2118  
Db 561 ProProAlaAlaAasnGlnLeuThrLeuSerAasnLeuHisLeuSerLeuSerLeu 580  
Qy 2119 TTAGCAAAACAATGCACTAGCATCTCTTACCAATCTCCAGCGCAAGATTCATCTCT 2178  
Db 581 LeuAlaAasnAasnAlaValThrAasnProProThrAasnProProAlaGlnAasnSerHisPro 600  
Qy 2179 CGACTCATTTGGTAGCAACTGCTGCTTCTGTTTACAAATTAGTGGGCTATCTTTTGTAG 2238  
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620

Qy 2239 GATTTCGATGATACAGCTTATGATAGGTATGATAGGTCTAGGTCTTAATCAAAAAATCAAT 2298  
Db 621 AspLeuAaspThrAlaTyrAaspAlaGlyTyrAaspTyrLeuGlySerAasnGlnIysIleAasp 640  
Qy 2299 GTCTCTGAAATACAGTTAGGACTAAGCCCCAGCTAAGTCCCAATGCCCCATCAGATTTGACTCTA 2358  
Db 641 ValLeuIysLeuGlnLeuGlyThrGlnProSerAlaAasnAlaProSerAaspLeuThrLeu 660  
Qy 2359 GCGAATGAGATGCTTAAGTATGCTATCAAGGAAGCTGGAAGCTTCGCTGGGATTCCTAAT 2418  
Db 661 GlyAasnGlnMetProIysTyrGlyTyrGlnGlySerTyrIysLeuAlaTyrPheAaspProasn 680  
Qy 2419 ACAGCAAAATTAATGCTCTTATCTCTGAAAGCTACATGGAATAAACTGGGTATTAATCCT 2478  
Db 681 ThrAlaAasnAasnGlyProTyrThrLeuIysAlaThrTyrThrIysThrGlyTyrAasnPro 700  
Qy 2479 GGGCTGAGGAGCTAGCTTCTTGGTTCCAAAATAGTTATGGGATCCATTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAasnSerLeuTyrGlySerIleLeuAaspIle 720  
Qy 2539 CGATCTGCGCATTCAGCAATTCGAATTTCTCTATCATCACCGCATGCTTTAGGTCAGGATAT 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAaspGlyArgSerTyrCysArgGlyLeu 740  
Qy 2599 TGGTCTTCTGAGGTTTCGAATTTCTCTATCATCACCGCATGCTTTAGGTCAGGATAT 2658  
Db 741 TrrValSerGlyValSerAasnPhePheTyrHisAaspArgAaspAlaLeuGlyGlnGlyTyr 760  
Qy 2659 CGGTATATTAGTGGGGTTTATCTTTAGGAGCAACTCTCTACTTTGGATCATCATGTTT 2718  
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAasnSerTyrPheGlySerSerMetPhe 780  
Qy 2719 GGTCTACCATTTCCGAAGTATTTGGTAGATCTAAAGATATTAGTATGATGCTGCTTCCAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerIysAaspTyrValValCysArgSerAasn 800  
Qy 2779 CATCATGCTGCTGATAGGATCCGTTTATCTATCTACCAACCAAGCTTTATGATGATCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrIysGlnAlaLeuCysGlySerTyr 820  
Qy 2839 TTGTTTCGAGATGCTGTTATCCGTCGCTAGCTACGGTTTGGGAATCAGCATATCAAAACC 2898  
Db 821 LeuPheGlyAaspAlaPheLeuArgAlaSerTyrGlyPheGlyAasnGlnHisMetIysThr 840  
Qy 2899 TCATATACATTTGCAGAGGAGCGATGCTCGTTGGGATAATAACTGCTGCTGCTGGAGAG 2958  
Db 841 SerTyrThrPheAlaGluGluSerAaspValArgTyrPheAasnAasnCysLeuValGlyGlu 860  
Qy 2959 ATTGAGCGGATTTACCGATTTGATTAATCTCACTCAATCTAAGCTCTATTTGAATGAGTTGCGT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerIysLeuTyrLeuAasnGluLeuArg 880  
Qy 3019 CTTTTCGTCGAACTGAGTTTCTTATCCCGATCATGATCTTTTACAGAGGAAGCGCAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAaspHisGluSerPheThrGluGluGlyAasp 900  
Qy 3079 CAAGCTCGGCAATTCACAGCGGACATCTCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaAlaGalaPheArgSerGlyHisLeuMetAasnLeuSerValProValGlyValIys 920  
Qy 3139 TTTGATTCGATGTTCTAGTACACATCTTAATAATATAGCTTTATCGCGCTTATATCTGT 3198  
Db 921 PheAaspArgCysSerSerThrHisProAasnIysTyrSerPheMetGlyAlaTyrIleCys 940  
Qy 3199 GATGCTTATCGCAACATCTCTGTTACTGAGACAAAGCTCTCTATCCATCAACAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACACAGAGATGCTTTTCATTTAGCAAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3318  
Db 961 ThrThrAaspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980



Db 381 AlalysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
Qy 1579 QATGTCGAGCGGATTTATTTAGGAAATCTCGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
Qy 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAT 1698  
Db 421 AspilellePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
Qy 1699 GCGTAACTGTGCTCTCAAGCCATTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyValIleThrLeu 460  
Qy 1759 AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818  
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
Qy 1819 AACCGCAGCGCAGCTCTCCAACTCTTAAATAATTAACGATGGTGAAGGATACACAGGG 1878  
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500  
Qy 1879 GATATTGTTTTGCTAAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGAGCAAGGA 1938  
Db 501 AspilelleValPheAlaAsnGlyAsnSerLeuTyrGlnAsnValThrIleGluGlnGly 520  
Qy 1939 AGGATTGTTCTTCGTAAGGCAAAATTTATCAGTGAATTCCTCTAAGTCAGACAGGTGG 1998  
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
Qy 1999 AGTCTGTATATGAAGCTGGGAGTACATGGGATTTGTAACTCCCAACACCACCAACAG 2058  
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
Qy 2059 CCTCTCGCGCTAAATCAGTTGATCAGCTTCCCAATCTGCAATTCCTCTCTCTCTCTG 2118  
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
Qy 2119 TTAGCAAAACAATGCAAGTATGATAGGTATGATGCTAGGTCTTAATCAAAAAATCAAT 2178  
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
Qy 2179 GAGTCATTTGGTAGCAACACTGCTGCTGTTCTGTACAAATTAGTGGGCTATCTTTTGG 2238  
Db 601 AlaIlelleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
Qy 2239 GATTTGGATGATACAGCTTATGATAGGTATGATGCTAGGTCTTAATCAAAAAATCAAT 2298  
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrPleuGlySerAsnGlnLysIleAsp 640  
Qy 2299 GTCTCGAAATTACAGTTAGGACTAAGCCCCAGCTAATGCCCCATCAGATTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
Qy 2359 GGAATAGATGCTTAAGTATGGCTATCAAGAGCTGGAAGCTTCGCTGGGATCCTAAT 2418  
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrPleuLysLeuAlaTyrAspProAsn 680  
Qy 2419 ACAGCAAAATAATGGTCTTATCTCTGAAAGCTACATGACTAAAGCTGGGTATATCTCT 2478  
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrIleThrLysThrGlyTyrAsnPro 700  
Qy 2479 GGGCTCGAGGAGTACTCTTTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspile 720  
Qy 2539 CGATCTGCCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGAGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
Qy 2599 TGGGTTCTCGAGTTTCGAATTTCTCTATCATGACCGGATGCTTTAGTCTCAGGATAT 2658

Db 741 TrpValSerGlyValSerAsnPheTyrHisPargAspAlaLeuGlyGlnGlyTyr 760  
Qy 2659 CCGTATATTATGTTGGGTTATTCCTTAGGACAAATCCTCTACTTTTGGATCATCGATGTT 2718  
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
Qy 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTGTTCCAAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
Qy 2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCCAAACAAAGCTTTATGGATCCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
Qy 2839 TTGTTCCGAGATGCGTTTATCCGCTAGCTACCGGTTTGGGAATCAGCATATGAAAACC 2898  
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
Qy 2899 TCATATACATTTGCAGAGGAGGAGCATGCTCGTTGGGATAATACTGCTCGCTCGAGAG 2958  
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860  
Qy 2959 ATTGAGCGGATTAACCGATTGTGATTACTCCATCTAAGCTCTATTGTAATGAGTTCGT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
Qy 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCCGATCATGAAATCTTTTACAGAGGAAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy 3079 CAAGCTCGGCATTCACAGAGCGGACATCTCTAAATCTATCATGTTCCCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
Qy 3139 TTTGATCATGTTCTAGTACATCATCTAATAATAATATAGCTTTATGGCGGCTTATATCTGT 3198  
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
Qy 3199 GATGCTTATCGCACCATCTCTGGTACTGAGCAACAGCTCCCTATCCCATCAGAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATCGAGTTGTGGTTAGAGGATCTATGATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
Qy 3319 TCTCTAAACAAGTAATATAGAGTATATGGCCATGGAAGATATGAGTATCGAGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
Qy 3379 CGAGCTATGTTTGGTTCGAGTCAGGAAGTAGAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

## RESULT 19

AEAL9000

ID AEAL9000 standard; protein; 1013 AA.

XX AC

XX AEAL9000;

XX DT 28-JUL-2005 (first entry)

XX DE Chlamydia trachomatis protein - SEQ ID 3.

XX KW chlamydia trachomatis infection; antibacterial; vaccine.

XX OS Chlamydia trachomatis.

XX PN US2005106162-A1.

XX PD 19-MAY-2005.

XX PP 16-DEC-2004; 2004US-00498327.



QY 1939 AGGATTGTTCTTCGTAAGGCAAAATTATCAGTGAATTTCTTAAGTCACAGAGTGGG 1998  
Db |||||  
521 ArgilleValLeuArgGluValAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
QY 1999 AGTCGTATATGGAGCTGGAGTACATGGGATTTTGTAACTCCACACCAACACAG 2058  
Db |||||  
541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560  
QY 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCATTTGCTCTTCTTCTTTG 2118  
Db |||||  
561 PropProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
QY 2119 TTAGCAACAATGCAAGTTAGCAATCCTCTCAATCTCCAGCCCAAGATTTCTCATCCT 2178  
Db |||||  
581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
QY 2179 GCAGTCATTTGGTAGCACAACTGCTGGTTCTGTTCATTAATTAGTGGCCCTATCTTTTGG 2238  
Db |||||  
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298  
Db |||||  
621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsp 640  
QY 2299 GTCTCTGAATTTACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA 2358  
Db |||||  
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
QY 2359 GGAATGAGATGCTCAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGATCCTAAT 2418  
Db |||||  
661 GlyAsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
QY 2419 ACAGCAAAATAATGTCCTTATCTCTGAAAGTACATGAGCTAAACCTGGGTAAATCT 2478  
Db |||||  
681 ThrAlaAsnAsnGlyProTyThrThrLeuLysAlaThrTrpThrLysThrGlyTyAsnPro 700  
QY 2479 GGGCTGAGCGAGTACCTTCTTGGTCCAAATAGTTTATGGGATCCATTTTACATATA 2538  
Db |||||  
701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGTCGAGGATTA 2598  
Db |||||  
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740  
QY 2599 TGGGTTCTGGAGTTTCGAATTTCTCTATCAGACCGCGATGCTTTAGTTCAGGATAT 2658  
Db |||||  
741 TrpValSerGlyValSerAsnPheTyThrHisAspArgAspAlaLeuGlyGlnGlyTy 760  
QY 2659 CGGTATATAGTGGGGTTATTCCTTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718  
Db |||||  
761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAsnSerTyPheGlySerSerMetPhe 780  
QY 2719 GGTCTAGCATTTACCGAAGTATTTGCTAGATCTTAAAGATTATGATGTGCTTCCAAT 2778  
Db |||||  
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAsn 800  
QY 2779 CATCATGCTTGATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTGGATCCTAT 2838  
Db |||||  
801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820  
QY 2839 TTGTTCCGGAGATCGGTTTATCGTCTAGCTAGCTAGGTTTGGGAATCAGCATATGAACC 2898  
Db |||||  
821 LeuPheGlyAspAlaPheLeuArgAlaSerTyGlyPheGlyAsnGlnHisMetLysThr 840  
QY 2899 TCATATACATTTGCAGAGAGAGCGATGTTCTGTGGATAATAACTGCTGCTGGAGAG 2958  
Db |||||  
841 SerTyThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
QY 2959 ATTGGACGGGATACCGATTGTGATTAATCTCCATCTTAAGCTCTTATTTGAATGAGTGGCT 3018  
Db |||||  
861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyLeuAsnGluLeuArg 880  
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGAT 3078

Db ||||| PropPheValGlnAlaGluPheSerTyArgAlaAspHisGluSerPheThrGluGluGlyAsp 900  
QY 3079 CAAGCTCGGCGATTCACAGAGCGGACATCTCCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138  
Db ||||| GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
QY 3139 TTTGATCGATGCTCTGATACACATCCCTAAATAATATAGCTTTATGGCGCTTATATCTGT 3198  
Db ||||| PheAspArgCysSerSerThrHisProAsnLysTySerPheMetGlyAlaTyIleCys 940  
QY 3199 GATGCTTATCCGACCATCTCTGGTACTGAGACAAAGCTCCTTATCCATCAAGAGACATGG 3258  
Db ||||| AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTCTTCAATTAGCAAGACATGAGTGTGGTGTAGAGGATCTATGATGCT 3318  
Db ||||| ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyAla 980  
QY 3319 TCTCTCAACAACTAATATAGATATATGGCCATCGAGATATGAGTATCGAGATGCTTCT 3378  
Db ||||| SerLeuThrSerAsnIleGluValTyGlyHisGlyArgTyGluTyArgAspThrSer 1000  
QY 3379 CGAGCTATGCTTTGAGTGCAGGAAGTACAGTCCGGTTC 3417  
Db ||||| ArgGlyTyGlyLeuSerAlaGlySerLysValArgPhe 1013  
RESULT 20  
AEA18998  
ID AEA18998 standard; protein; 1013 AA.  
XX AEA18998;  
AC AEA18998;  
XX  
DT 28-JUL-2005 (first entry)  
XX Chlamydia trachomatis protein - SEQ ID 1.  
XX Chlamydia trachomatis infection; antibacterial; vaccine.  
XX Chlamydia trachomatis.  
XX US2005106162-A1.  
XX 19-MAY-2005.  
XX 16-DEC-2004; 2004US-00498327.  
XX 12-DEC-2001; 2001GB-00029732.  
PR 06-AUG-2002; 2002EP-00182233.  
PR 14-AUG-2002; 2001EP-00218924.  
PR 12-DEC-2002; 2002WO-IB005761.  
XX (GRAN/) GRANDI G.  
PA (RATT/) RATTI G.  
XX Grandi G, Ratti G;  
PI WPI; 2005-354777/36.  
DR N-PSDB; AEA18999.  
XX  
PT New immunogenic composition comprising Chlamydia trachomatis protein or  
its encoding nucleic acid, useful for immunizing against, or treating or  
preventing chlamydial infection.  
XX Claim 13; SEQ ID NO 1; 90pp; English.  
XX  
CC The invention comprises an immunogenic composition for the treatment or  
prevention of a Chlamydia trachomatis infection. The immunogenic  
composition contains a Chlamydia trachomatis protein or its encoding  
nucleic acid. The immunogenic composition of the invention is useful for  
treating or preventing chlamydial infection. The present amino acid  
sequence represents a Chlamydia trachomatis protein of the invention.  
CC  
NOTS: The present sequence is not shown in the specification, but was



CC obtained from the USPTO website -  
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.  
 XX  
 SQ Sequence 1013 AA;

Alignment Scores:  
 Pred. No.: 0 Length: 1013  
 Score: 5121.50 Matches: 978  
 Percent Similarity: 98.8% Conservative: 23  
 Best Local Similarity: 96.5% Mismatches: 11  
 Query Match: 65.0% Indels: 1  
 DB: 9 Gaps: 1

US-10-701-844-1 (1-4435) x AEA18998 (1-1013)

QY 382 ATGCAACGCTTTCCATAGTCTTTCTTCAATGATTTAGCTTATTTCTGCTCT 441  
 DB 1 MetGlnThrSerPheHisLysPheLeuSerMetLeuLeuAlaTyrSerCysSer 20  
 QY 442 TTAATGGGGGATATGACGAGAAATCATGTTCTCCTCAGGAATTTACGATGGGAG 501  
 DB 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
 QY 502 ACCTTAACGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 561  
 DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrValPhe 60  
 QY 562 TCTGCAGGAGATTAACTTAAATAATCTTGACAATTTCTATGCACTTTCAGCTTTTAA 621  
 DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80  
 QY 622 TGTTTTGGGACCTTATAGGAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 681  
 DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
 QY 682 GAGAACATACGGACTTCTCAAAATGGGGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741  
 DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerSerAlaAsnSerGlyLeu 120  
 QY 742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTTCCAAATTCGAATTCATTACTTGC 801  
 DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnGlyAsnSerLeuLeuAlaVal 140  
 QY 802 CTCCTGCTGCACGACTAATAAGGTAGCCAGACTCCGACGACCAACATCTACACGCTCT 861  
 DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160  
 QY 862 AATGCTACTATTATTCTTAAACAGATCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 921  
 DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180  
 QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTTAAGCTTCAA 981  
 DB 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
 QY 982 GGAATTAGCAAGCTTTGCTTCTTCCAAAGAAATCTCTGCTCAAGCTGATGGGGAGCTTGT 1041  
 DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
 QY 1042 CAACTAGTACCAAGTTTCTGCTATGCTGCTAAGCGAGGCTCTTCTTCTTCTTCTTCTTCT 1101  
 DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
 QY 1102 GTTGCAGGATTAAGAGGGGAGGATTTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1161  
 DB 241 ValAlaGlyValArgGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
 QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTCTTCCAGAAATCTGCGGTAGAGTTT 1221  
 DB 261 SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
 QY 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTTCTTCTTCTTCTTCTTCTTCTG 1281

DB 281 AspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
 QY 1282 AATAATGGAACACCTTTGTTCTCAACATGTTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1341  
 DB 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrLeuAlaAlaGlu 320  
 QY 1342 CAACCAACAAAGTGGAGCGCTTCTTAATACGAGTAATAATATTCGAGAGATCGAGAGCTATC 1401  
 DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
 QY 1402 TTTCTGTAGAATGGTGGCAA---GCAGGATCCAATACTCTGATCGATGCTTCTTCTTCTTCT 1458  
 DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
 QY 1459 CGAGAGGAGTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1518  
 DB 361 GlyGluGlyValPhePheSerSerAsnValAlaAlaGlyGlyGlyGlyAlaIleTyr 380  
 QY 1519 GCCAAAAAGCTCTCGGTTCTTAACTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1578  
 DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
 QY 1579 GATGTGGAGCGATTTATTTAGAGATCTCGAGAGCTCAGTTTATCTGCTGATTTATGGA 1638  
 DB 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
 QY 1639 GATATTTATTTCTGATGGGAATCTTAAAGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1698  
 DB 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
 QY 1699 GCGCTAATCTGTCTCAACAGCCATTTCTGATGGGATCGGAGGGGAGGAGGAGGAGGAGGAG 1758  
 DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
 QY 1759 AGAGCTAAAGCAGGCGATCAGATTTCTTTTAAATGATCCCATCGAGATGGCAAGCAAGCA 1818  
 DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
 QY 1819 AACCAGCAGCGAGCTTCCAACTTCTTAAATTAAGCAAGTGTGAGGATGAGGATGAGGATGAG 1878  
 DB 481 AsnGlnProAlaGlnSerGluProLeuLysLysAlaAsnAspGlyGluGlyTyrThrGly 500  
 QY 1879 GATATTTGTTTCTTAAATGAGCAGTACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1938  
 DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
 QY 1939 AGGATTTGTTCTTCTGTAAGGCAAAATTTATCAGTGAATTTCTTAAAGTACAGAGGTGGG 1998  
 DB 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
 QY 1999 AGTCTGTATATGGAAGCTGGAGTACATGGGATTTTGTAACTCCACACCAACCAACACAG 2058  
 DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
 QY 2059 CCTCTGCGCTTAACTCAGTTGATCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2118  
 DB 561 ProProAlaAlaAsnGlnIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580  
 QY 2119 TTAGCAAAACAATGCAGTTAGCAATCTCTCTTACCAATCTCTCCAGCGCAAGATTTCTCATCT 2178  
 DB 581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
 QY 2179 GCAGTCAATGGTAGCAACTGCTGTTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2238  
 DB 601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGly 620  
 QY 2239 GATTTGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298  
 DB 621 AspLeuAspAspThrAlaTyrAspArgTyrAspThrLeuGlySerAsnGlnLysIleAsp 640  
 QY 2299 GTCTGTAATTTACAGTTAGGAGCTTAAGCCCCAGCTAATCCCCCATCAGATTTGACTCTA 2358  
 DB 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660

```
QY 2359 GGGAAATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGATCCTAAT 2418
Db 661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAen 680
QY 2419 ACAGCAATAATGGTCCCTTATCTACTGTAAGACTACATGAGCTAAAGCTGGGTATAATCCT 2478
Db 681 ThrAlaAenAenglyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAenPro 700
QY 2479 GGGCTGAGCAGTAGTCTCTTTGGTTCCTCAATAGTCTTATGGGATCCATTTTATGATATA 2538
Db 701 GlyProGluAArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGCGCAATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGTCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTCTATCATGACCGGATGCTTATAGGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAenPheTyrHisAspAArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTATTCTCTTAGGACAAACTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCAATTCAGCAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800
QY 2779 CATCATGCTGCATAGGATCGGTTTATCTATCTACCAACAGCTTATGTGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTTCGGAGATGCGTTTATCGTCTAGCTAGCTAGCGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheLeuargAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGCGATGTTCTGTTGGATAATACTGTCTGCTGGAGAG 2958
Db 841 SerTyrThrPheAlaGluSerAspValAlaGTrpAspAenAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGATTCAGGATTTGATTTACTCTTACCTTATGCTTATTTGATGAGTGGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGGCAATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGCTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTATAATATAGTCTTTATGGCGGCTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGACCATCTCTGCTACTGAGACAACTCTCTATCCATCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGACAGACATGAGGATTTGGTTAGAGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAATATAGAGTATATGCGCATGGAAGATATGATGATCGATGCTTCT 3378
Db 981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGGTTGAGTCAGGAAGTAGAGTCGGGTTCC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

## RESULT 21

```
AEA19014
ID AEA19014 standard; protein; 1013 AA.
XX AC AEA19014;
XX 28-JUL-2005 (first entry)
XX Chlamydia trachomatis protein - SEQ ID 17.
XX Chlamydia trachomatis infection; antibacterial; vaccine.
XX Chlamydia trachomatis.
XX US2005106162-A1.
XX 19-MAY-2005.
XX 16-DEC-2004; 2004US-00498327.
XX 12-DEC-2001; 2001GB-00029732.
XX 06-AUG-2002; 2002EP-00182233.
XX 14-AUG-2002; 2001EP-00218924.
XX 12-DEC-2002; 2002WO-IB005761.
XX (GRAN/) GRANDI G.
XX (RATT/) RATTI G.
XX Grandi G, Ratti G;
XX WPI; 2005-354777/36.
XX N-P9DB; AEA19015.
XX New immunogenic composition comprising Chlamydia trachomatis protein or
its encoding nucleic acid, useful for immunizing against, or treating or
preventing chlamydial infection.
XX Claim 13; SEQ ID NO 17; 90pp; English.
XX The invention comprises an immunogenic composition for the treatment or
prevention of a Chlamydia trachomatis infection. The immunogenic
composition contains a Chlamydia trachomatis protein or its encoding
nucleic acid. The immunogenic composition of the invention is useful for
treating or preventing chlamydial infection. The present amino acid
sequence represents a Chlamydia trachomatis protein of the invention.
NOTE: The present sequence is not shown in the specification, but was
obtained from the USPTO website -
seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX SQ Sequence 1013 AA;
```

```
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5112.50 Matches: 977
Percent Similarity: 98.7% Conservative: 23
Best Local Similarity: 96.4% Mismatches: 12
Query Match: 64.9% Indels: 1
DB: 9 Gaps: 1
```

US-10-701-844-1 (1-4435) x AEA19014 (1-1013)

```
QY 382 ATGCAAGCTCTTTCCATAAGTCTTTCTTCAATGATCTAGCTATTCTTCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetLeuLeuAlaTyrSerCysSer 20
QY 442 TTAATATGGGGGGATATGCACAGAAATCATGTTTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACCTTAACCTGATCATTTCCCTATATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
```

QY	562	TCGAGGAGAGTTAAACATTAATAAAATCTTGACAAATCTTATTCGAGCTTTGCTTTAAGT	621	Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
Db	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaAlaLeuProLeuSer	80	QY	1699	GGCGTAACGTGCTCTCAAGCCATTTTCGATGGGATCGGAGGAGAAATAACGACATTA	1758
QY	622	TCGTTTGGAACTTATTAGGAGTTTACTGTTTATAGGAGGAGCACTCGTTGACTTTTC	681	Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyPheLysIleThrLeu	460
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100	QY	1759	AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGCCAAACGGAAT	1818
QY	682	GAGACATACGACTCTCAAAATGGGCGAGCTCTAAGTAATAGCGCTCTGATGACATG	741	Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120	QY	1819	AACCGCAGCCGAGCTTCCAAACTTCTAAAAATTAACGATGCTCAAGATACACAGGG	1878
QY	742	TTTACTATTGAGGTTTTAAAGAAATATCTCTTTCCAAATTGCAATTCATTACTTGCCTA	801	Db	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyThrGly	500
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140	QY	1879	GATATTGTTTTCTTAATCGAAGCAGTACTTTGTACCAAAATGTTACGATAGACGAGGA	1938
QY	802	CTGCTGTGCAACGACTAATAGGCTAGCCAGACTCCGACGACCAACATCTACCGTCT	861	Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuThrGlnAsnValThrIleGluGlnGly	520
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer	160	QY	1939	AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGTGGG	1998
QY	862	AATGCTACTATTATTCTAAACAGATCTTTTGTACTCAATAAGAGAGTTCTCATTC	921	Db	521	ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
Db	161	AsnGlyThrIleThrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe	180	QY	1999	ACTCTGTATATGGAAGCTGGAGTACATGGGATTTTGTAACTTCTTAAGTCAGACAGTGGG	2058
QY	922	TATAGTAATTTAGTCTCTGAGATGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA	981	Db	541	SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln	200	QY	2059	CCTCTGCGCTAATCAGTTGATCAGCTTCCCAATCTGCATTTGCTCTCTCTCTCTTCT	2118
QY	982	GGAATTAGCAAGCTTTGTCTTCCAGAAATPACTGCTCAAGCTGATGGGAGCTTGT	1041	Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580
Db	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220	QY	2119	TTAGCAAAACAAATCGAGTTACGAATCTCTCTACCAATCTCCAGCGCAAGATTTCTATCCT	2178
QY	1042	CAAGTAGTCACCAAGTTCTCTGCTATGCTTAAGAGCTCTTATTCCTTCTGAGCAAT	1101	Db	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240	QY	2179	GCAGTCATTTGGTAGCAACACTGCTGTTCTGTTACAATTAGTGGGCTATCTTTTGTAG	2238
QY	1102	GTTCAGGAGTAGAGGGGAGGATTCGCTGTTCCAGATGCGCAGCAGGAGGTGCA	1161	Db	601	AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
Db	241	ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlyValSer	260	QY	2239	GATTGGATGATACAGCTTATGATAGTATGATGGCTAGTCTTCTAATCAAAAATCAAT	2298
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTCAGAAATATCTCGGTAGGTTT	1221	Db	621	AspLeuAspAspThrAlaTyArgTyArgTyArgTyArgTyArgTyArgTyArgTyArg	640
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280	QY	2299	GTCTCAAAATTAACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA	2358
QY	1222	GATGGAACTACCGGAGTAGAGGAGGATTTCTCTACCGGAGGAGCTGCTTCTCCTG	1281	Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTySerTyArgTyArgTyArgTyArgTyArg	300	QY	2359	GGGAATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCTTAAT	2418
QY	1282	ATAATGGAAACCTGTTTCTCAACAATGTTGCTTCTCTCTTACATTTGCTGCTAAG	1341	Db	661	GlyAsnGluMetProLysTyArgTyArgTyArgTyArgTyArgTyArgTyArgTyArg	680
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyLeuAlaAlaGlu	320	QY	2419	ACAGCAAAATAGTCTTATCTCTGAAAGCTACATGAGCTAATAAAGCTGGTATATCCT	2478
QY	1342	CAACCAACAGATGAGCGGCTCTTAATACGAGTAATAATTAACGAGATGAGGAGCTATC	1401	Db	681	ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTyThrLysThrGlyTyArgPro	700
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyArgTyArgTyArgTyArg	340	QY	2479	GGGCTGAGCGAGTACTCTTTGGTTTCCAAATAGTTTATGGGATCCATTTTAGATATA	2538
QY	1402	TTCTGTAAGAAATGGTGGCAAA--G--GAGGATCCAAATTAACCTGATGATGCTTCTT	1458	Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTyArgTyArgTyArgTyArg	720
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360	QY	2539	CGATCTCGCATTCAGCAATTCAGCAAGTGTGGTGGGCTCTTATTTGTCAGGATTA	2598
QY	1459	GGAGGGGAGTCTTTTCTTAGTACCAATGATGCTGCTGCGAAGGGGAGCTATTAT	1518	Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyArgTyArgTyArg	740
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTy	380	QY	2599	TGGGTTCTGGAGTTTCCGAATTTCTCTATCATGACCGCATCTTTAGCTCAGGATAT	2658
QY	1519	GCCAAAGCTCTCGGTGCTTAACCTGGGCCCTGTACAAATTTTAAAGAAATATCGCTAAT	1578	Db	741	TrpValSerGlyValSerAsnPheTyThrHisAspArgAspAlaLeuGlyGlnGlyTy	760
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400	QY	2659	CGGTATATTAGTGGGCTTATCTCTTAGGAGCAAACTCTCTACTTTGATCATCGATCTTT	2718
QY	1579	GATGGTGAGCGATTTATTTAGGAGAACTGGAGAGCTCAGTATTCTGCTGATTTAGGA	1638	Db	761	ArgTyThrIleSerGlyGlyTyThrLeuGlyAlaAsnSerTyThrPheGlySerSerMetPhe	780
Db	401	AspGlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyArg	420	QY	2719	GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTTATGTAGTGTCTGTTCAAT	2778
QY	1639	GATATTATTTCGATGGGAATCTTAAAGAACAGCAAGATGCTGCGGATGTTAAT	1698				

```
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy 2779 CATCATGCTTCATAGAGTCCGTTTATCTATCTATCCCAACAGCTTTTATGTGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTCCGAGATCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAACACC 2898
Db 821 LeupheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy 2899 TCATATACATTGCGAGAGCGCATGTTGTTGGGATATACATCTGCTGCTGGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860
Qy 2959 ATTGAGCGGATACCATGTTGATTAATCTCCATCAAGCTCTATTTCAATGAGTTGGGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy 3019 CTTTTCGTCGAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGCATTCACAGAGCGGCATCTCCTAAATCTATCAGTTCCCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTTGATCATGTTCTAGTACACATCCTTAATAAATAGCTTTTATGGCGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACACAGCTCCTATCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy 3259 ACAACAGATGCTTTTCATTAGCAGACATGAGTGTGTTAGAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Qy 3319 TCTCTAACAGTAATATAGAATATATGCCATGGAAGATATGAGTATCGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy 3379 CGAGCTATGTTTGAAGTCAGGAAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

## RESULT 22

AAV16737  
ID AAV16737 standard; protein; 1013 AA.

XX AC AAV16737;

XX DT 21-JUL-1999 (first entry)

XX DE C. trachomatis B serovar HMW protein.

XX KW Chlamydia; high molecular weight protein; HMW protein; urethritis;  
KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;  
KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;  
KW salpingitis; tubal occlusion; infertility; cervical cancer;  
KW arteriosclerosis; atherosclerosis.

XX OS Chlamydia trachomatis.

XX PN W09917741-A1.

XX PD 15-APR-1999.

XX PF 01-OCT-1998; 98MO-US020737.

XX PR 02-OCT-1997; 97US-00942596.

XX PA (ANTE-) ANTEK BIOLOGICS INC.

```
XX FI Jackson JW, Pace JL;  
XX DR WPI; 1999-287659/24.  
XX PT New Chlamydia protein useful for treating conjunctivitis, urethritis and  
XX cervical cancer.  
XX PS Claim 4; Page 115-119; 141pp; English.  
XX CC The invention relates to an isolated Chlamydia species high molecular  
CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as  
CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can  
CC be used for preventing, treating or ameliorating a disorder related to  
CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,  
CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,  
CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical  
CC cancer, infertility, arteriosclerosis and atherosclerosis. The products  
CC can also be used for detection and diagnosis. The present sequence  
CC represents a C. trachomatis HMW protein  
XX SQ Sequence 1013 AA;
```

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5106.50 Matches: 980  
Percent Similarity: 98.4% Conservative: 17  
Best Local Similarity: 96.7% Mismatches: 15  
Query Match: 64.8% Indels: 1  
DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x AAY16737 (1-1013)

```
Qy 382 ATGCAAAAGCTCTTCCATAAGTCTCTTCTTCAATGATTCTAGCTTATTCTGTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20  
Qy 442 TTAATGGGGGGGATATGACAGCAATCATGTTCTCTCAAGGAATTTACATGGGAG 501  
Db 21 LeuAsnGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40  
Qy 502 ACGTTAAGTGTATCATTTCCCTATCTATGAGGATCCGAGTCCGAGTCCGAGTCTGTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTGAGGAGAGTTAAACATTAAAAAATCTTGACAAATCTTATTCAGCTTTTGCCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
Qy 622 TGTGTTGGAACTTATTAGGAGTTTACTGTGTTTAGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy 682 GAGACATACCGACTTCTACAANTGGGCGACTCTAAGTAATAGCGCTGCTGATCGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
Qy 742 TTTACTATTGAGGGTTTTAAAGAATTATCTTTTCCAAATTCGAATTCATTTACTTCCCGTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140  
Qy 802 CTGCTGCTGCAACGACTTAATAAGGGTAGCCAGACTCCGACGACAACTCTACCGCTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrSerThrProSer 160  
Qy 862 AATGTTACTATTATTCTTAACACAGATCTTTTGTGTTACTCAATAATAGAGAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe 180  
Qy 922 TATAGTAATTTTAGTCTCTGGAGATCGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA 981  
Db 181 TyrSerAsnSerValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
```

QY 982 GGAATTACAGCTTTGCTCTTCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
DB |||||  
QY 201 GlyIleSerLysLeuCyvalPheGlnAenThrAlaGlnAlaAspGlyAlaCys 220  
DB |||||  
QY 1042 CAAGTAGTCACCAAGTTTCTCTGCTATGGCTAACAGAGCTCTTATTCCTCTTACCGAAT 1101  
DB |||||  
QY 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheValAlaAen 240  
DB |||||  
QY 1102 GTTCAGGAGTAGAGGGAGGAGTGTCTGCTTTCAGGATGGCGCAGGAGGAGTGCA 1161  
DB |||||  
QY 241 ValAlaGlyIleValArgGlyGlyIleAlaAlaValGlnAenGlyGlnGlnGlyValSer 260  
DB |||||  
QY 1162 TCATCTACTTCAACAGACAGCAGTAGTAGTGTTCCTCAGAAATACCTCGGTAGAGTTT 1221  
DB |||||  
QY 261 SerSerThrSerThrGlnAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
DB |||||  
QY 1222 GATGGGAACGTACCCGAGTAGGAGGAGGATTTACTCTACGGGAACCTGCTTTCCTG 1281  
DB |||||  
QY 281 AspGlyAenValAlaArgValGlyGlyIleIleTyrSerTyrGlyAenValAlaPheLeu 300  
DB |||||  
QY 1282 AATATGGAAAAACCTGTTTCTCAACATGTTGCTTCTCCTGTTTACATTGCTGCTAAG 1341  
DB |||||  
QY 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaAlaGlu 320  
DB |||||  
QY 1342 CAACCAACAGTGGCAGGCTTCTAATACAGTAATAATTTACGAGATGGAGGAGCTATC 1401  
DB |||||  
QY 321 GlnProThrAenGlyGlnAlaSerAenThrSerAspAsnTyrGlyAspGlyAlaIle 340  
DB |||||  
QY 1402 TTCTGTGAAGATGGTCCGCAA---GCAGGATCCAAATAACTCTGGATCAGTTTCTTTGAT 1458  
DB |||||  
QY 341 PheCyLeuAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAsp 360  
DB |||||  
QY 1459 GGAGGAGGAGTAGTTTCTTTTAGTACAAATGATGCTGCGGAAAGGGGAGCTATTAT 1518  
DB |||||  
QY 361 GlyGluGlyValValPhePheSerAenValAlaAlaGlyLysGlyAlaIleTyr 380  
DB |||||  
QY 1519 GCCTAAAGCTCTCGGTCTTAACCTGCGCCCTGATACAAATTTTAAAGGAATATCGCTAAT 1578  
DB |||||  
QY 381 AlaLysLysLeuSerValAlaAenCyGlyProValGlnLeuLeuGlyAenIleAlaAen 400  
DB |||||  
QY 1579 GATGTGGAGCGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
DB |||||  
QY 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
DB |||||  
QY 1639 GATATTATTTTCATCGGAACTTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAT 1698  
DB |||||  
QY 421 AspMetIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440  
DB |||||  
QY 1699 GCGCTAACTGTGCTCACAGCCATTTGATCGGATCGGAGGAGGAAATAACGACATTA 1758  
DB |||||  
QY 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIleThrThrLeu 460  
DB |||||  
QY 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTAATGATCCCATCGAGATGGCAACCGGAAAT 1818  
DB |||||  
QY 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAenGlyAen 480  
DB |||||  
QY 1819 AACCCAGCCAGCGAGTCTTCCAAACTTTCAAAATTTAAACCATGCTGAAGATACACAGG 1878  
DB |||||  
QY 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAenAspGlyGlyIleTyrThrGly 500  
DB |||||  
QY 1879 GATATTGTTTTGCTAATGAAGCAGTACTTTGTTACCAAAATGTTACGATAGGACGAGGA 1938  
DB |||||  
QY 501 AspIleValPheAlaAenGlyAenSerThrLeuTyrGlnAenValThrIleGluGlnGly 520  
DB |||||  
QY 1939 AGGATCTTCTCTGTAAGGCGCAAAATTTATCAGTGAATTTCTCTAAGTCAGACAGGTGGG 1998  
DB |||||  
QY 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540  
DB |||||  
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTCTAATCTCCAAACCAACCAACAG 2058  
DB |||||  
QY 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560  
DB |||||  
QY 2059 CCTCTCGCCGCTAATCAGTTGATCAGCTTTCCAAATCTGCAATTTGTTCTCTTCTTTG 2118  
DB |||||

DB |||||  
QY 561 ProProAlaAlaAenGlnSerIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeu 580  
DB |||||  
QY 2119 TTACCAACAATCGAGTTTACGAATCTCTACCAATCTCTCAGCGCAAGATTCTCATCT 2178  
DB |||||  
QY 581 LeuAlaAenAenAlaValThrAenProThrAenProProAlaGlnAspSerHisPro 600  
DB |||||  
QY 2179 GCAGTCATTTGGTAGCACCAACTGCTGTTCTGTACAAATTTAGTGGCCCTATCTTTTGG 2238  
DB |||||  
QY 601 AlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
DB |||||  
QY 2239 GATTTGGATGATACAGCTTATGATAGGTATGATTCGCTAGGTCTTCAATCAAAATCAAT 2298  
DB |||||  
QY 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAenGlnLysIleAsp 640  
DB |||||  
QY 2299 GTTCTGAAATTTACAGTTAGGACTAAGCCCCAGCTAATGCCCAATCAGATTTCATCTTA 2358  
DB |||||  
QY 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAenAlaProSerAspLeuThrLeu 660  
DB |||||  
QY 2359 GCGAATGAGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCGGATCCTAAT 2418  
DB |||||  
QY 661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTyrLeuAlaTrpAspProAen 680  
DB |||||  
QY 2419 ACAGCAAAATAATGGTCTTATCTCTGAAAGCTACATGCACTAAACTGGGTATATCTCT 2478  
DB |||||  
QY 681 ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAenPro 700  
DB |||||  
QY 2479 GGGCTGAGCGAGTAGTCTTCTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538  
DB |||||  
QY 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTyrGlySerIleLeuAspIle 720  
DB |||||  
QY 2539 CCATCTGCGCATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAAT 2598  
DB |||||  
QY 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCyArgGlyLeu 740  
DB |||||  
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTTATCATATGACCCGAGATGCTTTAGTCTAGGATAT 2658  
DB |||||  
QY 741 TrpValSerGlyValSerAenPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
DB |||||  
QY 2659 CGGTATATTTAGTGGGGTTTATCTCTTAGGAGCAAACTCTTACTTTCGATCATCGATGTTT 2718  
DB |||||  
QY 761 ArgTyrIleSerGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780  
DB |||||  
QY 2719 GGTCTAGCATTTACCAAGTATTTGGTAGATCTTAAAGATTTAGTGTGCTGCTCCAAT 2778  
DB |||||  
QY 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCyArgSerAen 800  
DB |||||  
QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTACCCAAACAGCTTTTATGTCGATCTAT 2838  
DB |||||  
QY 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
DB |||||  
QY 2839 TTGTTCCGAGATCGCTTTTATCCGTCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898  
DB |||||  
QY 821 ValPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840  
DB |||||  
QY 2899 TCATATACATTTTCAGCAGGAGCGATGTTTGGTGGGTAATAACTCTCTGGCTGGAGAG 2958  
DB |||||  
QY 841 SerTyrThrPheAlaGluLeuSerAspValCysTrpAspAenCysLeuValGlyGlu 860  
DB |||||  
QY 2959 ATTGGAGCGGATTTACCGATTGTCGATTTACTTCCATCTTAAGCTCTTATTTGAATGTCGCT 3018  
DB |||||  
QY 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880  
DB |||||  
QY 3019 CTTTTCGTGCAAGCTCAGATTCTTATGCGCATCATGATCTTTTACAGAGGAGGCGAT 3078  
DB |||||  
QY 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGlySerPheThrGluGluGlyAsp 900  
DB |||||  
QY 3079 CAGCTCGGCAATTCAGAGCGGACATCTCTAAATCTATCAATCTCTGTTGAGTCAAG 3138  
DB |||||  
QY 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920  
DB |||||  
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATTAATATAGCTTTTATGGCGGCTTATCTCT 3198  
DB |||||

Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
 QY 3199 GATGCTTATCGCACCATTCTCTGGTACTGAGACAACGCTCCTATCCATCAAGAGACATGG 3258  
 Db 941 AspaLalTyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960  
 QY 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGAGAGTTGGTTAGAGGATCTATGATGCT 3318  
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
 QY 3319 TCTCTAACAGTAAATATAGATATATGCGCATGGAAGATATGATCGAGATGCTTCT 3378  
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
 QY 3379 CGAGGCTATGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417  
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 23

ID AAV16738 standard; protein; 1013 AA.

AC AAV16738;

DT 21-JUL-1999 (first entry)

DE C. trachomatis F serovar HMW protein.

KW Chlamydia; high molecular weight protein; HMW protein; urethritis;  
 KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;  
 KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;  
 KW salpingitis; tubal occlusion; infertility; cervical cancer;  
 KW arteriosclerosis; atherosclerosis.

OS Chlamydia trachomatis.

XX WO9917741-A1.

XX 15-APR-1999.

XX 01-OCT-1998; 98WO-US020737.

XX 02-OCT-1997; 97US-00942596.

XX (ANTE-) ANTEX BIOLOGICS INC.

XX Jackson JW, Pace JL;

XX WPI; 1999-287659/24.

XX New Chlamydia protein useful for treating conjunctivitis, urethritis and  
 PT cervical cancer.

XX Claim 4; Page 119-123, 141pp; English.

XX The invention relates to an isolated Chlamydia species high molecular  
 CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as  
 CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can  
 CC be used for preventing, treating or ameliorating a disorder related to  
 CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,  
 CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,  
 CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical  
 CC cancer, infertility, arteriosclerosis and atherosclerosis. The products  
 CC can also be used for detection and diagnosis. The present sequence  
 CC represents a C. trachomatis HMW protein

XX Sequence 1013 AA;

XX Alignment Scores:

Pred. No.:	0	Length:	1013
Score:	5106.50	Matches:	980
Percent Similarity:	98.2%	Conservative:	15
Best Local Similarity:	96.7%	Mismatches:	17

Query Match:	64.8%	Indels:	1
DB:	2	Gaps:	1
US-10-701-844-1 (1-4435) x AAY16738 (1-1013)			
QY	382	ATGCAAAAGCTTTCCATTAAGTCTCTTCAATGATTTCTAGCTTATTTCTTGCTGCTCT	441
Db	1	MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer	20
QY	442	TTAAATGGGGGGGATATGACAGCAAAATCATGTTCTCAAGGAATTTACGATGGGAG	501
Db	21	LeuThrGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu	40
QY	502	ACGTTAACTGATCATTTCCCTATCTATTAGAGATCCGAGTGGGACTACTGTTTTT	561
Db	41	ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTGCAGGAGAGTTAAACATTAATAAATCTTCACAAATCTTATTCGACGTTTGCCTTAA	621
Db	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
QY	622	TGTTTGGGAACTTATTAGGGAGTTTACTGTTTGGGAGAGGACACTCGTTGACTTTC	681
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
QY	682	GAGAACATACGAGCTTCTCAAAATGGGGAGCTCTAAGTAATAGCGCTGCTGATGACTG	741
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
QY	742	TTTACTATTGAGGGTTTAAAGATTATCCCTTTCCAAATTCGAATTCATTTACTTCGCTA	801
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGCTGCAACGACTAATAAGGCTAGCAGACTCCGACGACAAACATCTACACCGCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	AATGTTACTATTATTCTTAAACAGACTTTTGTGTTTCTCAATTAATGAGAGTTCTCATTC	921
Db	161	AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlyLysPheSerPhe	180
QY	922	TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGCTGATCGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyThrIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAATTAGCAAGCTTTGTGTTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTCT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys	220
QY	1042	CAAGTAGTCACAGTTTCTCTGCTATGCTAACGAGGCTCTTATTCCTTTGCTAGGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTCGAGGATGAGAGGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1161
Db	241	ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTGCGGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGAACTGACCGGAGTAGGAGGAGGATTTACTCTTACCGGAACTGCTGCTTCTCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAAACCTTTGTTCTCAACATGTTGCTTCTCTGTTTACATTCGCTGCTAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu	320
QY	1342	CAACCAACAGATGGAGCGCTTCTAATACGACTAATTAATACGAGATGGAGGAGTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle	340





KW Chlamydia; vaccine; infection; fusion protein; antigen;  
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;  
 KW acute respiratory tract infection; Cap; CT529; OMCB; TSA.  
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.  
 XX Chlamydia trachomatis.  
 XX W02001040474-A2.  
 XX 07-JUN-2001.  
 XX 04-DEC-2000; 2000WO-US032919.  
 XX 03-DEC-1999; 99US-00454684.  
 XX 19-APR-2000; 2000US-00556877.  
 XX 20-JUN-2000; 2000US-00598419.  
 XX (CORI-) CORIXA CORP.  
 XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;  
 XX WPI; 2001-374831/39.  
 XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic  
 XX inflammatory disease, trachoma, acute respiratory tract infections,  
 XX atherosclerosis and heart disease.  
 XX Claim 2; Page 214-216; 295pp; English.  
 XX The present sequence is provided in a specification relating to compounds  
 XX and methods for the treatment and diagnosis of chlamydial infection. The  
 XX compounds provided include polypeptides and fusion proteins comprising  
 XX immunogenic portions of Chlamydia antigens and DNA sequences encoding  
 XX such polypeptides. They are useful for vaccinating against chlamydial  
 XX infection, which causes pelvic inflammatory disease, trachoma, acute  
 XX respiratory tract infections, atherosclerosis and heart disease  
 XX Sequence 1006 AA;

Alignment Scores:  
 Pred. No.: 0 Length: 1006  
 Score: 5090.00 Matches: 979  
 Percent Similarity: 99.9% Conservative: 2  
 Best Local Similarity: 99.7% Mismatches: 1  
 Query Match: 64.6% Indels: 0  
 DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x AAG83207 (1-1006)

QY 472 ATGGTTCCTCAAGGAATTTACGATGGGAGAGCTTAACGTATCATTTCCCTATACTGTT 531  
 DB 25 MetileProGInGlyIleYrAspGlyGluThrLeuThrValSerPheProfyThrVal 44  
 QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTCTCGCAGGAGAGTTAAACATTTAAAAATCTT 591  
 DB 45 lIeGlyAppProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64  
 QY 592 GACAAATCTATTGACAGCTTTCCTTTAAGTTCTTTGGGAATTAATAGGGAGTTTACT 651  
 DB 65 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84  
 QY 652 GTTTTAGGAGGAGGACCTGTTGACTTTCGAGACATACGACTTCTACAAATGGGCA 711  
 DB 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104  
 QY 712 GCTCTAAGTAATAGCGCTGCTGATGACTGTTTACTATTAGAGGTTTAAAGAAATATCC 771  
 DB 105 AlaLeuSerAsnSerAlaAlaaspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124  
 QY 772 TTTTCCAATTCGAATTCATTACTGCGGTATGCGCTGCTGCAACGACTAATAGGGTAC 831  
 DB 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144

QY 832 CAGACTCCGACGACAACTCTACCGTCTAATGTTACTATTTATTTCTAAACAGATCTT 891  
 DB 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleYrSerLysThrAspLeu 164  
 QY 892 TTGTTACTCAATAATCAGAAGTTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGGA 951  
 DB 165 LeuLeuLeuAsnAsnGluLysPheSerPheYrSerAsnLeuValSerGlyAspGlyGly 184  
 QY 952 GCTATAGATGCTAAGAGCTTAACGGTTTCAAGGAATTAGCAAGCTTTGTGTTCTCAAGAA 1011  
 DB 185 AlaIleaspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 204  
 QY 1012 AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCT 1071  
 DB 205 AsnThrAlaGlnAlaaspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224  
 QY 1072 AACGAGCTCTTATGCTTTGTAGGAATGTTGCAGGAGTAAAGAGGGGAGGATGCT 1131  
 DB 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 244  
 QY 1132 GCTGTTCCAGGATGGCAGCAGGAGTGTCTATCTACTTCAACAGAGATCCAGTAGTA 1191  
 DB 245 AlaValGlnaspGlyGlnGlnGlyValSerSerThrThrThrGluAspProValVal 264  
 QY 1192 AGTTTTTCCAGAAATACTGCGTAGAGTTTGTATGGGAACGTAGCCCGAGTAGGAGGG 1251  
 DB 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaAa-gValGlyGlyGly 284  
 QY 1252 ATTTACTCTACGGAACTGCTTCTCTGAATATCGMAAAACCTGTTTCTCAACAAAT 1311  
 DB 285 IleYrSerYrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 304  
 QY 1312 GTTCTCTCTCTGTTTACATTTGCTTAAGCAACCAACAGTGACAGGCTTCTAATACG 1371  
 DB 305 ValAlaSerProValYrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324  
 QY 1372 AGTAATAATTACGGAGATGGAGGAGCTATCTTCTGTAAAGATGTTGCGCAACGAGGATCC 1431  
 DB 325 SerAsnAsnYrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344  
 QY 1432 AATAACTCTGATCAGTTTCTCTGATGGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
 DB 345 AsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSerAsnVal 364  
 QY 1492 GCTGCTGGGAAAGGGAGCTATTTATGCCAAAAGCTCTCGGTTGCTTAACGTGCGCCT 1551  
 DB 365 AlaAlaGlyLysGlyGlyAlaIleYrAlaLysLysLeuSerValAlaAsnCysGlyPro 384  
 QY 1552 GTACAAATTTTAAAGGAATATCGTAATCATGTTGGTGGAGCGATTTATTTAGGAGAATCTGA 1611  
 DB 385 ValGlnPheLeuAaGAsnIleAlaAsnAspGlyGlyAlaIleYrLeuGlyGlySerGly 404  
 QY 1612 GAGCTCAGTTTATCTGCTGATTTGGAGATATTATTTTCGATGGGAATCTTAAAGAAACA 1671  
 DB 405 GluLeuSerLeuSerAlaAspYrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424  
 QY 1672 GCCAAGAGAAATGCTGCCGATTAATGCGGTAACTGTGCTCTCAAGACCAATTTCCATG 1731  
 DB 425 AlaLysGluAsnAlaAlaaspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444  
 QY 1732 GGATCGGAGGGAATAATACGACATTAAGAGCTAAAGCGGCGCATCAGATCTCTTATAT 1791  
 DB 445 GlySerGlyGlyLysIleThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464  
 QY 1792 GATCCCTCAGATGGCAAAACGGAATAACACGAGCGAGCTTCTCCAACTTCTTAAAA 1851  
 DB 465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484  
 QY 1852 ATTTACGATGGTGAAGGATACACAGGGATATTGTTTTTGTCTAATGAAGCAGTACTTGT 1911  
 DB 485 IleAsnAspGlyGlyGlyYrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504  
 QY 1912 TACCAAAATGTTACGATAGCAAGGAGGATTTGTTCTGCGTGAAGGCAAAATTTATCA 1971

Db	505	TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLeuYsaAlaYsaLeuSer	524	865	SerLysLeuTyrLeuAsnGluLeuArgPropheValGlnAlaGluPheSerTyrAlaAep	884
QY	1972	GTGAATCTCTAAGTCAGACAGGTGGAGTCTGTATATAGGAAGCTGGAGTACATGGGAT	2031	3052	CATGAATCTTTTACAGAGAGAGCCATCAAGCTCGGGCATTCAGAGCGGACATCTCCTA	3111
Db	525	ValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThrLeuAep	544	885	HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu	904
QY	2032	TTTGTAACCTCCACACACACACAGCTCTCGCGCTAAATCAGTTGATCAGCTTCC	2091	3112	AATCTATCAGTTCTGTGGAGTGAAGTTGATCGATGTTCTAGTACACATCCTAATAA	3171
Db	545	PheValThrProGlnProGlnProGlnProAlaAlaAsnGlnLeuIleThrLeuSer	564	905	AnLeuSerValProValGlyValLysPheAepArgCysSerSerThrHisProAsnLys	924
QY	2092	AATCTGCAATTTGCTCTTCTCTTTGTTAGCAAAACAATGCGATTACGAATCCTCTACC	2151	3172	TATAGCTTTATGGCGGCTTATATCTGTCATGCTTATCGCACCCTCTCTGTTACTGAGACA	3231
Db	565	AnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr	584	925	TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr	944
QY	2152	AATCTCCAGCGCAAGATCTCATCTCGCAGTCATTGGTAGCAAACTGCTGTTCTGTT	2211	3232	ACGCTCTATCCCATCAAGAGACATGGCAACAGATGCTCTTTCATTTAGCAAGACATGGA	3291
Db	585	AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrAlaGlySerVal	604	945	ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly	964
QY	2212	ACATTAGTGGGCTATCTTTTGGAGATTTGGATGATACAGCTTATGATAGGTATGAT	2271	3292	GTTGTGGTTAGAGGATCTATGCTTCTCTTCTTAAACAAGTAATATAGAGTATATGGCCAT	3351
Db	605	ThrIleSerGlyProIlePhePheGluAepLeuAspAspThrAlaTyrAspArgTyrAep	624	965	ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis	984
QY	2272	TGGCTAGGTTCTAATCAAAAAATCAATGCTCTGAAATACAGTTAGGGACTAAGCCCCA	2331	3352	GGAAATATAGATATCGAGATGCTCTCGAGGCTATGTTTTCAGTGCAGAGTAGAGTC	3411
Db	625	TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro	644	985	GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal	1004
QY	2332	GCTAATGCCCATCAGATTGACTTAGGGAATGAGATGCTAAGTAGGCTATCAAGGA	2391	3412	CGGTTC	3417
Db	645	AlaAsnAlaProSerAepLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly	664	1005	ArgPhe	1006
QY	2392	AGCTGAAGCTTCGCTGGGATCTTAATACAGCAAAATAGTCTTATCTCTGAAGCT	2451	RESULT 25		
Db	665	SerTrpLysLeuAlaTrpAepProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla	684	ID	ABB94178	standard; protein; 1006 AA.
QY	2452	ACATGACTAAACTGGGTATATCTTGGGCTCGAGGAGTAGCTTCTTGGTTCCAAAT	2511	XX	ABB94178;	
Db	685	ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn	704	DT	05-JUN-2002	(first entry)
QY	2512	AGTTATGGGGATCCATTTTATAGATACGATCGCGCATTCAGCAATTCAGCAAGTGTG	2571	DE	Chlamydia protein sequence SEQ ID NO:190.	
Db	705	SerLeuTrpGlySerIleLeuAepIleArgSerAlaHisSerAlaIleGlnAlaSerVal	724	XX	Chlamydia infection; Chlamydia; vaccine; detection; diagnosis; antigen;	
QY	2572	GATGGGGCTCTTATGTCAGGATATATGGTTCTGGAGTTTCGAATTTCTTCTATCAT	2631	KW	antibacterial; immunostimulant; immune response;	
Db	725	AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis	744	KW	Chlamydia-specific T-cell response.	
QY	2632	GACCGGATCTTTAGTCAAGGATATCGGTATATTAGTGGGGTATTCTTCTAGGAGCA	2691	OS	Chlamydia sp.	
Db	745	AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla	764	PN	W0200208267-A2.	
QY	2692	AATCTCTACTTTGGATCATCGATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT	2751	XX	31-JAN-2002.	
Db	765	AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer	784	XX	20-JUL-2001; 2001WO-US023121.	
QY	2752	AAAGATATATAGTGTGTTCCATCATCATCTGTCATGATCCGTTTATCTATCT	2811	XX	20-JUL-2000; 2000US-00620412.	
Db	785	LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer	804	PR	23-APR-2001; 2001US-00841132.	
QY	2812	ACCCACAAGCTTTATGTGATCTTATTTGTCGGAGATCGTTTATCCGTGCTAGCTAC	2871	XX	(CORI-) CORIXA CORP.	
Db	805	ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr	824	XX	Fling SP, Skeiky YAW, Probst P, Bhatia A;	
QY	2872	GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAGCGCATGTTCTGT	2931	XX	WPI; 2002-179901/23.	
Db	825	GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluSerAspValArg	844	XX	Novel compositions comprising Chlamydia Cap1 protein and its use in the	
QY	2932	TGGGATATATCTGCTGGCTGGAGAGATTGGAGCGGGATTACCGATTGATTAATCCCA	2991	PT	treatment of Chlamydia infection.	
Db	845	TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro	864	XX	Example 1; Page 245-247; 537pp; English.	
QY	2992	TCTAAGCTCTATTGAATGAGTTGCGTCTTCTGTCGCAAGCTAGTTTCTTATGCGCAT	3051	CC	The present invention describes compositions comprising a Chlamydia Cap1	
				CC	protein and methods for the diagnosis and therapy of Chlamydia infection.	
				CC	Chlamydia DNA and protein sequences from the present invention can have	
				CC	antibacterial and immunostimulant activities, and can be used in	
				CC	CC vaccines. Compounds from the present invention can be used for eliciting	
				CC	an immune response, specifically stimulating a Chlamydia-specific T-cell	
				CC	response or inhibiting the development of a Chlamydia infection in an	

CC animal. Methods from the present invention can be used: for detecting the  
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells  
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia  
 CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent  
 CC sequences used in the exemplification of the present invention

XX Sequence 1006 AA;

SQ Alignment Scores:

Pred. No.: 0 Length: 1006  
 Score: 5090.00 Matches: 979  
 Percent Similarity: 99.9% Conservative: 2  
 Best Local Similarity: 99.7% Mismatches: 1  
 Query Match: 64.6% Indels: 0  
 DB: 5 Gaps: 0

US-10-701-844-1 (1-4435) x ABB94178 (1-1006)

QY	472	ATGTTCTCTCAAGGAATTACGATGGGAGAGCTTAACGTATCATCTTCCCTACTACTGT	531
DB	25	MetIleProGInGlyIleTyAspGlyGluThrLeuThrValSerPheProTyThrVal	44
QY	532	ATAGAGAGATCCGAGTGGGACTACTCTTTTCTGCGAGAGAGTTAAACATTAAAAATCTT	591
DB	45	IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLeuAsnLeu	64
QY	592	GACAACTTCTATTGCGAGCTTGCTTTTAAGTTGTTTGGAACTTATTAGGAGTTTACT	651
DB	65	AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr	84
QY	652	GTTTTAGGAGAGGACACTCGTTGACTCTTCGAGAACATACGAGCTTCTACAAATGGGCA	711
DB	85	ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla	104
QY	712	GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGGTTTAAAGAATTATCC	771
DB	105	AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGlyPheGlyGluLeuSer	124
QY	772	TTTTTCCAATTGCAATTCATTACTTCCGCTACTGCTCTGCAACGACTTAATAAGGGTAGC	831
DB	125	PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer	144
QY	832	CAGACTCGACGACACATCTACCGCTTAATGTTACTATTATTCTTAAACAGACTCTT	891
DB	145	GlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLysThrAspLeu	164
QY	892	TTGTACTCAATAATAGAGATTCTCATTTAGTAATTTAGTCTCTGGAGATGGGA	951
DB	165	LeuLeuLeuAsnAsnGlnLysPheSerPheTySerAsnLeuValSerGlyAspGlyGly	184
QY	952	GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCAAGAA	1011
DB	185	AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu	204
QY	1012	AATAGCTCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCT	1071
DB	205	AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla	224
QY	1072	PACGAGGCTCTATTGCTTTCTGCGAATGTTGCGAGGATGAGAGGGGAGGAGTTGCT	1131
DB	225	AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla	244
QY	1132	GCTGTTTCAGGATGGGCGAGGAGTGTCTATCTTACTTCAACAGAAGATCCAGTAGTA	1191
DB	245	AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal	264
QY	1192	AGTTTTCCAGAAATCTACGGTAGAGTTTGATGGGAAGCTAGCCCGAGTAGGAGAGGG	1251
DB	265	SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly	284
QY	1252	ATTTACTCTACGGGAACGCTGCTTTCCTGAATATATGGAAGAACCTTGTCTCAACAA	1311
DB	285	IleTySerTyGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn	304

QY	1312	GTTCGCTTCTCTGTTTACATTGCTGTAAAGCAACCAAGATGGACAGGCTTCTTAATACG	1371
DB	305	ValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr	324
QY	1372	AGTAATAATTACGGAGATGGAGGAGCTATCTTCTGTAAGATGTCGCGACAGGAGATCC	1431
DB	325	SerAsnAsnTyGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer	344
QY	1432	AATAACTCTGATCAGTTTCTCTTGTGAGGAGGAGTAGTTTCTTTTAGTAGCAATGTA	1491
DB	345	AsnAsnSerGlySerValSerPheAspGlyGlyValValPhePheSerSerAsnVal	364
QY	1492	GCTGCTGGAAAGGGGAGACTATTATGCCCCAAAAGCTCTCGGTGCTTAACCTGGCCCT	1551
DB	365	AlaAlaGlyLysGlyGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCysGlyPro	384
QY	1552	GTACAAATTTTAAAGAAATATCGCTAATCATGCTGGAGCGATTTATTTAGGAGAATCTGA	1611
DB	385	ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyLeuGlyGluSerGly	404
QY	1612	GAGCTCAGTTTATCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTTAAAGAACA	1671
DB	405	GluLeuSerLeuSerAlaAspTyGlyAspIlePheAspGlyAsnLeuLysArgThr	424
QY	1672	GCCAAAGAGAAATGTCGCGATGTTAATGGCGTAACTGTGTCTCTCAAGCCATTTCGATG	1731
DB	425	AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet	444
QY	1732	GGATCGGAGGAAATTAACGACATTAAGACTAAAGCAGGCGCATCAGATCTCTTTAAT	1791
DB	445	GlySerGlyGlyLysIleThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn	464
QY	1792	GATCCCATCGAGATGGCAACCGAAATAACGAGCAGCGAGTCTTCCAAACTCTTAAAA	1851
DB	465	AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys	484
QY	1852	ATTAAAGATGTGAAGGATACACAGGGGATATGTTTTGCTTAATGGAAGCAGACTTTG	1911
DB	485	IleAsnAspGlyGlyGlyTyThrGlyAspIleValPheAlaAsnGlySerSerThrLeu	504
QY	1912	TACCAAAATGTCAGATAGCAAGCAAGGAGGATGTTCTTCGTGAAAGCGCAAAATATCA	1971
DB	505	TyrGlnAsnValThrIleGlnGlnGlyArgIleValLeuArgGlyLysAlaLysLeuSer	524
QY	1972	GTGAATCTCTAAGTCAGACAGGAGTCTGTATATGGAAGCTGGGAGTACATGGGAT	2031
DB	525	ValAsnSerLeuSerGlnThrGlySerLeuTyMetGluAlaGlySerThrLeuAsp	544
QY	2032	TTTGTAACTCCAAACCAACCAAGCCCTCTGCGCTTAATTCAGTTGATCAGCTTTCC	2091
DB	545	PheValThrProGlnProGlnGlnProAlaAlaAsnGlnLeuIleThrLeuSer	564
QY	2092	AATCTGCATTTGCTCTTCTTCTTGTAGCAACAAATGAGTTCACCACTCTCTCTACC	2151
DB	565	AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAlaValThrAsnProProThr	584
QY	2152	AATCTCCAGCGCAAGATTCTCATCTGCACTTGTGTAGCACAACCTGCTGTTCTGT	2211
DB	585	AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal	604
QY	2212	ACAATTAGTGGGCTTATCTTTTGTAGGATTTGGATGATACAGCTTATGATAGTATGAT	2271
DB	605	ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyAspArgTyAsp	624
QY	2272	TGGCTAGCTTCTTAATCAAAAATCAATCTCTGTAATTTACAGTTAGGAGCTTAAGCCCCA	2331
DB	625	TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro	644
QY	2332	GCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCCCTTAAGTATGGCTTATCAAGA	2391
DB	645	AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyGlyTyGlnGly	664

QY 2392 AGCTGGAGCTTGGCTGGATCCTAATACAGCAAAATATGGTCTTATCTCTGAAGACT 2451  
 Db |||||  
 QY 665 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 684  
 Db |||||  
 QY 2452 ACATGGACTAAAGCTGGGTATATCTCGGCTCGAGCGAGTAGCTTCTTGGTCCAAAT 2511  
 Db |||||  
 QY 685 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 704  
 Db |||||  
 QY 2512 AGTTATGGGATCCATTAGATATACGATCGGCATTCAGCAATTCAGCAAGTGG 2571  
 Db |||||  
 QY 705 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724  
 Db |||||  
 QY 2572 GATGGCGCTCTATTGTTCGAGATATGCGGTTCCTGAGTTCGAGATTTCTTCTATCAT 2631  
 Db |||||  
 QY 725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 744  
 Db |||||  
 QY 2632 GACCGCGATGCTTTAGTCAGGATATCGGTATATTAGTGGGGTATTCTTAGAGCA 2691  
 Db |||||  
 QY 745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764  
 Db |||||  
 QY 2692 AACTCTACTTTGGATCATCGATTTGGTCTAGCATTTACCGAGTATTGGTAGATCT 2751  
 Db |||||  
 QY 765 AsnSerTyrPheGlySerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784  
 Db |||||  
 QY 2752 AAGATTATGTAGTGTCTTCCCAATCATCATCTGCTTCATAGATCCGTTTATCTATCT 2811  
 Db |||||  
 QY 785 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 804  
 Db |||||  
 QY 2812 ACCCAACAAGCTTTATGTGGATCCTATTGTTCTCGAGATCGGTTTATCGGTGCTACTAC 2871  
 Db |||||  
 QY 805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824  
 Db |||||  
 QY 2872 GGTGTTGGGAATCAGCATATGAACCTCATATACATTTGTCAGAGAGCGATGTCGT 2931  
 Db |||||  
 QY 825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 844  
 Db |||||  
 QY 2932 TGGGTAATAACCTGCTGCTGGCTGAGAGATTGGAGCGGATTACCGATTGATTTACTCCA 2991  
 Db |||||  
 QY 845 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864  
 Db |||||  
 QY 2992 TCTAAGCTCTATTGTAATGAGTTGCGTCTCTTTCGTCAGAGCTGAGTTTCTTATGCCGAT 3051  
 Db |||||  
 QY 865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884  
 Db |||||  
 QY 3052 CATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCCTTA 3111  
 Db |||||  
 QY 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904  
 Db |||||  
 QY 3112 AATCTATCAGTTCTGTTGGAGTGAGTTGTGATCGATGTTCTAGTACACATCTCTAATAA 3171  
 Db |||||  
 QY 905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924  
 Db |||||  
 QY 3172 TATAGCTTTATGGCGCTTATCTGATGCTTATCGCCACTCTCTGCTACTGAGACA 3231  
 Db |||||  
 QY 925 TyrSerPheMetAlaIleTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944  
 Db |||||  
 QY 3232 ACGCTCTATCCCATCAAGAGACATGACACAGATCCCTTCTATTAGCAAGACATGGA 3291  
 Db |||||  
 QY 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964  
 Db |||||  
 QY 3292 GTTGTGTTAGAGATCTATGATGCTTCTCAAGATTAATAGAGTATATGAGCAT 3351  
 Db |||||  
 QY 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984  
 Db |||||  
 QY 3352 GGAAGATATGATATCAGATGCTTCTCGAGGCTTATGTTGAGTCAGGAGTAGATC 3411  
 Db |||||  
 QY 985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004  
 Db |||||  
 QY 3412 CGGTTT 3417  
 Db |||||  
 QY 1005 ArgPhe 1006

RESULT 26

AA013633  
 ID AAB13633 standard; protein; 982 AA.  
 XX AAB13633;  
 AC AAB13633;  
 XX 02-FEB-2001 (first entry)  
 DT  
 XX C. trachomatis pmpG gene protein.  
 DE  
 XX Chlamydial infection; sexually transmitted disease;  
 KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
 KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;  
 KW coronary heart disease; antibacterial.  
 XX Chlamydia trachomatis.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 981  
 FT /note= "Unspecified amino acid"  
 XX  
 PN W0200034483-A2.  
 XX  
 PD 15-JUN-2000.  
 XX  
 PF 08-DEC-1999; 99WO-US029012.  
 XX  
 PR 08-DEC-1998; 98US-00208277.  
 PR 08-APR-1999; 98US-00288594.  
 PR 01-OCT-1999; 99US-00410568.  
 PR 22-OCT-1999; 99US-00426571.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
 XX  
 DR WPI; 2000-431303/37.  
 XX  
 PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
 PT comprises immunogenic portion of Chlamydia antigen, which comprises amino  
 PT acid sequence encoded by polynucleotide sequence.  
 XX  
 PS Claim 2; Page 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydiae are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention  
 XX  
 SQ Sequence 982 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 982  
 Score: 5084.00 Matches: 978  
 Percent Similarity: 99.8% Conservatives: 2  
 Best Local Similarity: 99.6% Mismatches: 2  
 Query Match: 64.5% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-701-844-1 (1-4435) x AAB13633 (1-982)  
 QY 472 ATGCTTCTCAAGAAATTTACGATGGGAGCGTAACTGATCATCTTCCTACTACTGTT 531  
 |||:|  
 Db 1 MetileProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20



Db 741 AnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760  
 Qy 2752 AAAGATTATGTAGTGTGCTTCCCAATCATATGCTTGGATAGGATCGTTTATCTATCT 2811  
 Db 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780  
 Qy 2812 ACCCAACAGCTTATGTGATCCTATTTGTTGGAGATGCGTTTATCGTGCTAGCTAC 2871  
 Db 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800  
 Qy 2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGCATGTTCGT 2931  
 Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820  
 Qy 2932 TGGGATAATAACTGCTCTGCTGGAGAGATGGAGCGGATATACCGATGTGTGATCTCCA 2991  
 Db 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840  
 Qy 2992 TCTAAGCTCTATTTGAATGATGTTGCTCTCTTCTGTCAGCTGAGTCTTCTTATGCGAT 3051  
 Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860  
 Qy 3052 CATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGCGATTCAGAGCGGACATCTCCTA 3111  
 Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880  
 Qy 3112 AATCTATCACTCTCTGTTGGAGTGAAGTGTGATCGATGTTCTAGTACACATCTATAATA 3171  
 Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900  
 Qy 3172 TATAGCTTATGCGGCTTATATCTGTGATGCTTATCGACCATCTCTGCTAGTACGACA 3231  
 Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920  
 Qy 3232 ACGCTCTATCCATCAAGAGACATGACACACAGATGCTTTCATTTAGCAAGACATGGA 3291  
 Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrAspAlaPheHisLeuAlaArgHisGly 940  
 Qy 3292 GTTGTGGTATAGAGATCTATGTATGCTTCTTAACAAGTAAATATAGATATATGCGCAT 3351  
 Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960  
 Qy 3352 GGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTGATGTCGAGGAGTATAGTC 3411  
 Db 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980  
 Qy 3412 CGGTTTC 3417  
 Db 981 \*\*\*Phe 982

RESULT 27

ID AAG83201 standard; protein; 982 AA.

XX AAG83201;

AC AAG83201;

DT 05-SEP-2001 (first entry)

XX Protein encoded by Chlamydia trachomatis pmpG gene.

XX Chlamydia; vaccine; infection; fusion protein; antigen;  
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;  
 KW acute respiratory tract infection; Cap1; CT529; OMCB;  
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.  
 XX Chlamydia trachomatis.

OS WO200140474-A2.

XX 07-JUN-2001.

XX 04-DEC-2000; 2000WO-US032919.

XX

PF

XX 03-DEC-1999; 99US-00454684.  
 PR 19-APR-2000; 2000US-00556877.  
 PR 20-JUN-2000; 2000US-00598419.  
 XX (CORI-) CORIXA CORP.  
 XX Probat P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;  
 XX WPI; 2001-374831/39.  
 XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic  
 PT inflammatory disease, trachoma, acute respiratory tract infections,  
 PT atherosclerosis and heart disease.  
 XX Claim 2; Page 189-191; 295pp; English.  
 XX The present sequence is provided in a specification relating to compounds  
 CC and methods for the treatment and diagnosis of chlamydial infection. The  
 CC compounds provided include polypeptides and fusion proteins comprising  
 CC immunogenic portions of Chlamydia antigens and DNA sequences encoding  
 CC such polypeptides. They are useful for vaccinating against chlamydial  
 CC infection, which causes pelvic inflammatory disease, trachoma, acute  
 CC respiratory tract infections, atherosclerosis and heart disease  
 XX Sequence 982 AA;  
 SQ

Alignment Scores:

Pred. No.: 0 Length: 982  
 Score: 5084.00 Matches: 978  
 Percent Similarity: 99.8% Conservative: 2  
 Best Local Similarity: 99.6% Mismatches: 2  
 Query Match: 64.5% Indels: 0  
 DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x AAG83201 (1-982)

Qy 472 ATGGTTCTCTCAAGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTACTGTT 531  
 Db 1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20  
 Qy 532 ATAGGAGATCCGAGTGGGACTACTGTTTTTCTGAGGAGAGTTAACTAAATAATCTT 591  
 Db 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40  
 Qy 592 GACAACTTATTCACGCTTTCCTTTAACTGTTGTTGGGAACCTATTAGGAGCTTTTACT 651  
 Db 41 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60  
 Qy 652 GTTTTAGGAGGAGGACACTCGTTGACTTTCGAGAACATACGAGCTTCTCAAAATGGGCA 711  
 Db 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80  
 Qy 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTATCTATTATGAGGTTTAAAGAAATATCC 771  
 Db 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 100  
 Qy 772 TTTTCCAAATGCAATTCATTACTTTCGCTACTGCTGCTGCTCAACGACTAATAGGCTAGC 831  
 Db 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 120  
 Qy 832 CAGACTCCGACGACACATCTACACGCTCTAAATGGTACTATTATTCTAAACACAGATCTT 891  
 Db 121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleThrSerLysThrAspLeu 140  
 Qy 892 TTGTTACTCAATAATGAGAAATGTTCTCATTTCTATAGTAATTTAGTCTCTGAGATGGGGA 951  
 Db 141 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160  
 Qy 952 GCTATAGATGCTAAGACCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAA 1011  
 Db 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180

QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTGTCTAAGTAGTCAACAGTTTCTCTGCTATGGCT 1071  
DB 181 AenThrAlaGlnAlaAaspGlyGlyAlaCysGlnValThrSerPheSerAlaMetAla 200  
QY 1072 AACGAGGCTCCTATTGCTTTGTAGCGGAATGTTGCAGGAGTAAGAGGGAGGAGTTCT 1131  
DB 201 AenGlnAlaProIleAlaPheValAlaAenValAlaGlyValArgGlyGlyIleAla 220  
QY 1132 GCTGTTACGATGGCAGCAGGAGTGTCTATCTACTTCAACAGAAGATCCAGTAGTA 1191  
DB 221 AlaValGlnAaspGlyGlnGlnGlyValSerSerSerThrThrGluAaspProValVal 240  
QY 1192 AGTTTTCAGAAATACCTGCGGTAGACTTTGATGGGAACGTAGCCGAGTAGGAGGG 1251  
DB 241 SerPheSerArgAenThrAlaValGluPheAaspGlyAenValAlaArgValGlyGly 260  
QY 1252 ATTTACTCTACGGGAAGCTTCTTCTGTAATATGGAATAACCTTGTCTTCAACAAT 1311  
DB 261 IleTyrSerTyrGlyAenValAlaPheLeuAenGlnGlyTyrLeuPheLeuAenAen 280  
QY 1312 GTTGCTTCTCTGTTTACATTTGCTCTAAGCAACCAACAGTGGACAGCTTCTAATACG 1371  
DB 281 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAenThr 300  
QY 1372 AGTAATAATTAGGAGATGGAGAGCTATCTTCTGTAAAGATGGTGGCAAGCAGGATCC 1431  
DB 301 SerAenAenTyrGlyAaspGlyAlaIlePheCysLysAenGlyAlaGlnAlaGlySer 320  
QY 1432 AATAACTCTGATCAGTTCTCTTTCATGGAGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
DB 321 AenAenSerGlySerValSerPheAaspGlyGluGlyValAlaPhePheSerSerAenVal 340  
QY 1492 GCTGCTGGAAAGGGGAGCTATTATGCAAAAAGCTCTCGTTGCTAACTGTGGCCCT 1551  
DB 341 AlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAenCysGlyPro 360  
QY 1552 GTACAAATTTTAAAGAAATATCCTTAATGATGTGGAGCGATTATTTAGGAGAATCTGA 1611  
DB 361 ValGlnPheLeuArgAenIleAlaAenAaspGlyAlaIleTyrLeuGlyGluSerGly 380  
QY 1612 GAGCTCAGTTTATCTGCTGATATCGAGATATTTTTCGATGGGAATCTTAAAGAAC 1671  
DB 381 GluLeuSerLeuSerAlaAaspTyrGlyAaspIleIlePheAaspGlyAenLeuLysArgThr 400  
QY 1672 GCCAAAGAGAAATGCTGCCGATGTTAATGGCGTAACCTGTCTCACAAGCCATTTGATG 1731  
DB 401 AlaLysGluAenAlaAaspValAenGlyValThrValSerSerGlnAlaIleSerMet 420  
QY 1732 GGATCGGAGGGAATAATACGACATTAGAGCTAAAGCGAGGCGATCAGATTCTCTTAAT 1791  
DB 421 GlySerGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAen 440  
QY 1792 GATCCCATCGAGATGGCAACGAAATAACACGACGCGAGCTTTCCAAACTTCTTAA 1851  
DB 441 AaspProIleGluMetAlaAenGlyAenAenGlnProAlaGlnSerSerLysLeuLeuLys 460  
QY 1852 ATTAACGATGGTGAAGGATACACAGGGGATATTTGTTTTGCTTAATGGAGCAGTACTTTG 1911  
DB 461 IleAenAaspGlyGluGlyTyrThrGlyAaspIleValPheAlaAenGlySerSerThrLeu 480  
QY 1912 TACCAAAATGTTACGATAGCAGGAGGAGGATTTGTTCTTGGTGAAGGCAAAATATCA 1971  
DB 481 TyrGlnAenValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500  
QY 1972 GTGAATTCCTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGGAGTACATGGGAT 2031  
DB 501 ValAenSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAasp 520  
QY 2032 TTTGTAACTCCAAACCAACCAACAGCCTCTCGCGCTTAATCAGTTGATCAGCTTTTCC 2091  
DB 521 PheValThrProGlnProGlnGlnProProAlaAlaAenGlnLeuIleThrLeuSer 540  
QY 2092 AATCTGCATTTGCTCTTCTTGTGTTAGCAACAAATGCAAGTACGAATCTCTCTACC 2151

DB 541 AenLeuHisLeuSerLeuSerSerLeuLeuAlaAenAenAlaValThrAenProProThr 560  
QY 2152 AATCTCTCAGCGCAGAGATTTCTCATCTCTGCACTTATGGTAGCACAACTGCTGTTCTGTT 2211  
DB 561 AenProProAlaGlnAaspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580  
QY 2212 ACAATTAGTGGGCTATCTTTTGGAGATTGGATGATATACAGCTTATGATAGTATGAT 2271  
DB 581 ThrIleSerGlyProIlePhePheGluAaspLeuAaspThrAlaTyrAaspArgTyrAasp 600  
QY 2272 TGGCTAGTCTTAATCAAAAATCAATGCTCTGCTCAAAATTTACAGTTTAGGAGCTAAAGCCCCA 2331  
DB 601 TrpLeuGlySerAenGlnLysIleAenValLeuLysLeuGlnLeuGlyThrLysProPro 620  
QY 2332 GCTAATGCCCATCAGATTTTGACTCTAGGGGAATGAGATGCTTAAGTATGGCTATCAAGA 2391  
DB 621 AlaAenAlaProSerAaspLeuThrLeuGlyAenGluMetProLysTyrGlyTyrGlnGly 640  
QY 2392 AGCTGGAAAGCTTGGTGGGATCTTAATACACAAATAATATGCTCTTATCTCTGAAAGCT 2451  
DB 641 SerTrpLysLeuAlaTrpAaspProAenThrAlaAenAenGlyProTyrThrLeuLysAla 660  
QY 2452 ACATGGCACTAAACTGGGTATAATCTGGGCTGAGCGAGTAGCTTCTTGGTTCCAAAT 2511  
DB 661 ThrTrpThrLysThrGlyTyrAenProGlyProGluArgValAlaSerLeuValProAen 680  
QY 2512 AGTTTATGGGATCCATTTTAGATATACGATCTCGCATTCGCAATTCAGCAATTCAGCAAGTGTG 2571  
DB 681 SerLeuTrpGlySerIleLeuAaspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700  
QY 2572 GATGGGCTCTTATGTGCGAGGATTAATGGGTTCTGGAGTTTCAATTTCTTCTATCAT 2631  
DB 701 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAenPhePheTyrHis 720  
QY 2632 GACCGCATGCTTAGGTCAGGATATCGGTATATTAGTGGGGTTATTCCTTAGGAGCA 2691  
DB 721 AaspArgAaspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740  
QY 2692 AACTCTACTTTGATCATCGATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751  
DB 741 AenSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760  
QY 2752 AAAGATTATGATGTGTGCTTCCATCATCATCATGCTTCCATAGGATCCGTTTATCTATCT 2811  
DB 761 LysAaspTyrValValCysArgSerAenHisAlaCysIleGlySerValTyrLeuSer 780  
QY 2812 ACCCAACAGCTTTATGAGTTCCTTATTTGTCGGAGATGCTTTATCCGTGCTAGCTAC 2871  
DB 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAaspAlaPheIleArgAlaSerTyr 800  
QY 2872 GGGTTTGGGAATCAGCATATGAAACCTCATATATATTTCAGAGAGAGCGATGTTCTG 2931  
DB 801 GlyPheGlyAenGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAaspValArg 820  
QY 2932 TGGGATAATAACTGCTGCTGGAGAGATTGGAGCGGATTAACCATTTGTTGTTACTTCCA 2991  
DB 821 TrpAaspAenAenCysLeuAlaGlyIleGlyAlaGlyLeuProIleValIleThrPro 840  
QY 2992 TCTAAGCTCTATTTCGAATGAGTTGCTCTTTCGTGCAAGCTGAGTTTCTTATCCGAT 3051  
DB 841 SerLysLeuTyrLeuAenGlnLeuArgProPheValGlnAlaGluPheSerTyrAlaAasp 860  
QY 3052 CATGAATCTTTTACAGAGAAAGCGATCAAGCTCGGGCATTCAGAGCGGACATCTCTTA 3111  
DB 861 HisGluSerPheThrGluGluGlyAaspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880  
QY 3112 AATCTATCAGTTCTGTTGGAGTGAAGTTTGATTCGATGTTCTAGTACACATCTCTATAAA 3171  
DB 881 AenLeuSerValProValGlyValLysPheAaspArgCysSerSerThrHisProAenLys 900  
QY 3172 TATAGCTTTATGGCGGCTTATATCTGATGCTTATCGCACCATCTCTGGTATCTGAGACA 3231



Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920  
 QY 3232 AGCTCTCATCCATCAAGACATGGACAAAGATGCTTTCATTTAGCAAGACATGGA 3291  
 Db 921 ThrLeuLeuSerHieGlnGluThrTrpThrThrAspAlaPheHieLeuAlaArgHieGly 940  
 QY 3292 GTTGTGGTTAGAGGATCTATGATGCTCTCTAAACAAGTAATATAGAGTATATGGCCAT 3351  
 Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHie 960  
 QY 3352 GGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTTCAGTGCAGGAAGTATAGTC 3411  
 Db 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerIysVal 980  
 QY 3412 CGGTTT 3417  
 Db 981 \*\*\*Phe 982

RESULT 28  
 ID ABB94172 standard; protein; 982 AA.  
 AC ABB94172;  
 XX 05-JUN-2002 (first entry)  
 DT XX  
 XX Chlamydia protein sequence SEQ ID NO:176.  
 DE XX  
 XX Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;  
 KW Chlamydial; immunostimulant; immune response;  
 KW Chlamydia-specific T-cell response.  
 XX Chlamydia sp.  
 OS XX  
 XX WO200208267-A2.  
 PN XX  
 XX 31-JAN-2002.  
 PD XX  
 XX 20-JUL-2001; 2001WO-US023121.  
 XX 20-JUL-2000; 2000US-00620412.  
 PR 23-APR-2001; 2001US-00841132.  
 XX (CORI-) CORIXA CORP.  
 PA Fling SP, Skeiky YAW, Probst P, Bhatia A;  
 PI WPI; 2002-179901/23.  
 XX Novel compositions comprising Chlamydia CapI protein and its use in the  
 PT treatment of Chlamydia infection.  
 XX Disclosure; Page 221-223; 537pp; English.

The present invention describes compositions comprising a Chlamydia CapI protein and methods for the diagnosis and therapy of Chlamydia infection. Chlamydia DNA and protein sequences from the present invention can have antibacterial and immunostimulant activities, and can be used in vaccines. Compounds from the present invention can be used for eliciting an immune response, specifically stimulating a Chlamydia-specific T-cell response or inhibiting the development of a Chlamydia infection in an animal. Methods from the present invention can be used: for detecting the presence of Chlamydia in a patient; to stimulate and/or expand T cells specific for a Chlamydia protein; and for treatment of a Chlamydia infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent sequences used in the exemplification of the present invention.

XX Sequence 982 AA;

Alignment Scores:  
 Pred. No.: 0 Length: 982  
 Score: 5084.00 Matches: 978  
 Percent Similarity: 99.8% Conservative: 2

Best Local Similarity: 99.6% Mismatches: 2  
 Query Match: 64.5% Indels: 0  
 DB: 5 Gaps: 0  
 US-10-701-844-1 (1-4435) x ABB94172 (1-982)  
 QY 472 ATGTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATCTACTGTT 531  
 Db 1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20  
 QY 532 ATAGAGATCCGAGTGGGACTACTGTTTCTTCGAGGAGATTAACTTAAATAATCTT 591  
 Db 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuIysAsnLeu 40  
 QY 592 GACAATTTCTATGACGCTTTTGGCTTTTAAAGTTCTTTTGGAACTTATAGGAGTTTACT 651  
 Db 41 LepAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60  
 QY 652 GTTTTAGGAGAGACACTCGTTGACTTTTCGAGACATACGAGCTTCTACAAATGGGGA 711  
 Db 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80  
 QY 712 GCTCTAAGTAATAGCGCTGCTGATCGACTGTTTACTATTGAGGGTTTAAAGAAATATCC 771  
 Db 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheIysGluLeuSer 100  
 QY 772 TTTTCCCAATTGCAATTCTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGCTAGC 831  
 Db 101 PheSerAsnSerCysAsnSerLeuLeuAlaValLeuProAlaIleThrThrAsnIysGlySer 120  
 QY 832 CAGACTCCGACGACCAACATCTACACCGTCTAATGGTACTATTATTCTTAAACAGATCTT 891  
 Db 121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerIysThrAspLeu 140  
 QY 892 TTGTTACTCAATAATGAGAGTTCATCTCTATAGTAATTTAGTCTCTCGAGATGGGGA 951  
 Db 141 LeuLeuAsnAsnGluIysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160  
 QY 952 GCTATAGATGCTAAGAGCTTAAACGGTTCAAGAAATTTAGCAAGCTTTGTCTTCCCAAGAA 1011  
 Db 161 AlaIleAspAlaIysSerLeuThrValGlnGlyIleSerIysLeuCysValPheGlnGlu 180  
 QY 1012 AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCCACAGTTTCTCTGCTATGGCT 1071  
 Db 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200  
 QY 1072 AACAGGCTCTATTGCTCTTTAGCGAATGTTGCGAGGATGAGAGGGGAGGATGCT 1131  
 Db 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 220  
 QY 1132 GCTGTTCCAGATGGGCGAGGAGTGTCATCTACTTCAACAGAAATCCAGTAGTA 1191  
 Db 221 AlaValGlnAspGlyGlnGlnGlyValSerSerThrSerThrGluAspProValVal 240  
 QY 1192 AGTTTTTCCAGAAATACTCGGCTAGATTGTTGAGAACGCTAGCCCGAGTAGGAGGG 1251  
 Db 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 260  
 QY 1252 ATTTACTCTACGGGAACGTTGCTTCTGTAATATGGAACAACTTCTTCAACAAAT 1311  
 Db 261 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyIysThrLeuPheLeuAsnAsn 280  
 QY 1312 GTTGCTTCTCTGTTTACATGCTGCTAAGCAACCAAGTCGACAGGCTTCTTAATACG 1371  
 Db 281 ValAlaSerProValTyrIleAlaAlaIysGlnProThrSerGlyGlnAlaSerAsnThr 300  
 QY 1372 AGTAATAATTACGAGATGGAGAGCTATCTTCTGTAAGAATGTCGCAAGCAGGATCC 1431  
 Db 301 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysIysAsnGlyAlaGlnAlaGlySer 320  
 QY 1432 AATAACTCTGGATCAGTTCTCTTTTGTAGGAGAGGAGTAGTTTCTTTTAGTACGAATGTA 1491  
 Db 321 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 340

QY 1492 GCTGCTGGGAAGGGGAGCTATTATATGCGCAAAAGCTCTCGTTGCTAACTGGCCCT 1551  
 Db 341 AlaAlaGlyLysGlyAlaAlaIleTyrAlaLysLeuSerValAlaAsnCyeglyPro 360  
 QY 1552 GTACAAATTTAAAGGAATATCGTAATGATGGTGGAGCGATTATTTAGGAGAACTCGA 1611  
 Db 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyAlaIleTyrLeuGlyGlnSerGly 380  
 QY 1612 GAGCTCAGTTTCTGCTGATTATCGAGATATATTTTCGATGGGAATCTTAAAGAACA 1671  
 Db 381 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400  
 QY 1672 GCCAAGAGAATGTCGCGATGTTAATGCGTAACCTGTGCTCAAGCCCAATTCGATG 1731  
 Db 401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420  
 QY 1732 CGATCGGAGGGAAAATACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTCTTAAT 1791  
 Db 421 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440  
 QY 1792 GATCCCATCGAGATGGCAACCGAAATAACAGCCAGCGCAGTCTTCCAAACTTCTAAAA 1851  
 Db 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys 460  
 QY 1852 ATTAACGATGGTGAAGGATACACAGGGATATTGTTTGTCTTAATGGAGCAATCTTG 1911  
 Db 461 IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480  
 QY 1912 TACCAAAATGTTACGATAGAGCAAGAGGATTCTTCGTGAAGAGCAAAATATCA 1971  
 Db 481 TyrGlnAsnValThrIleGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500  
 QY 1972 GTGAATCTCTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGAGTACATGGGAT 2031  
 Db 501 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 520  
 QY 2032 TTTGTAACCTCAACACCAACAGCCTCTCGCGCTTAATCAGTTGATCAGCTTCC 2091  
 Db 521 PheValThrProGlnProGlnProGlnProAlaAlaAsnGlnLeuIleThrLeuSer 540  
 QY 2092 AATCTGCAATGCTCTCTCTCTCTGTTAGCAAAATGCAATGCAATGCAATCCTCTACC 2151  
 Db 541 AsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560  
 QY 2152 AATCTCCAGCGCAAGATTCTCATCTGCAAGTCAATGGTAGCACAACTGCTGGTCTGTT 2211  
 Db 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrAlaGlySerVal 580  
 QY 2212 ACAATTAGTGGCCCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGGTATGAT 2271  
 Db 581 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 600  
 QY 2272 TGCTAGGTTCTAATCAAAAATCAATGTCCTGAAATACAGTTAGGTAAGGCCCA 2331  
 Db 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620  
 QY 2332 GCTAATGCCCCATCAGATTGACTCTAGGGAATGAGATGCTTAAGTATGCTATCAAGGA 2391  
 Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 640  
 QY 2392 AGCTGGAAGCTTGGCGGATCCTAATACAGCAAAATAAATGGTCTTATATCTCTGAAGCT 2451  
 Db 641 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660  
 QY 2452 ACATGCACTAAACTGGGTATATCTCGGCTGAGGAGTAGCTCTTTGGTTCCAAT 2511  
 Db 661 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680  
 QY 2512 AGTTTATGGGATCCCAATTTAGATATACGATCTGCGCATTCAGCAATTCAGCAAGTGTG 2571  
 Db 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700

QY 2572 GATGGCGCTTATTGTCGAGGATTATGGGTTTCTCGAGTTTCGAAATTTCTTCTATCAT 2631  
 Db 701 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720  
 QY 2632 GACCGCATGCTTTAGGTACGGATATCGGTATATTAGTGGGGTTATTCCTTAGGACA 2691  
 Db 721 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740  
 QY 2692 AACTCTACTTTGGATCATCATGATGTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751  
 Db 741 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760  
 QY 2752 AAAGATTATGATGTTGCTGCTTCCAAATCATCATGCTGCATAGGATCCGTTTATCTATCT 2811  
 Db 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleIleGlySerValTyrLeuSer 780  
 QY 2812 ACCCAACAAGCTTTATGTGATCTTATTTGTTGCGAGATGCGTTTATCCGTGCTAGCTAC 2871  
 Db 781 ThrGlnGlnAlaLeuCyseGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800  
 QY 2872 GGGTTTGGGAATCAGCATATCAAAACCTCATATACATTTGCGAGGAGGAGCATGTTCTGT 2931  
 Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820  
 QY 2932 TGGGATAATAACTGCTGCTGCGAGATTCGAGCGGATACCGGATTACCGATTGTGATTACTCA 2991  
 Db 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840  
 QY 2992 TCTAAGCTCTATTGAAATGAGTTCGCTCTTTCGCAAGCTGAGTTTCTTATCCGAT 3051  
 Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860  
 QY 3052 CATGAATCTTTTACAGAGGAAGGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCCTA 3111  
 Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880  
 QY 3112 AATCTCATGTTCTGTTGGAGTGAAGTTTCATCGATGCTTCATGACATCCCTAATAAAA 3171  
 Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900  
 QY 3172 TATAGCTTTATGGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231  
 Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920  
 QY 3232 AGCTCTCTATCCCATCAAGAGACATGAGACACAGATGCTTTCATTTAGCAAGACATGGA 3291  
 Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940  
 QY 3292 GTTGTGGTTAGGAGTCTATGATGCTTCTCTAACAAGTAATATAAGATATATGGCCAT 3351  
 Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960  
 QY 3352 GGAAGATATGATATCGAGATGCTTCTCGAGCTATGTTGAGTGCAGGAAAGTAGAGTC 3411  
 Db 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980  
 QY 3412 CGGTTTC 3417  
 Db 981 \*\*\*Phe 982  
 RESULT 29  
 AAB13639  
 ID AAB13639 standard; protein; 1006 AA.  
 XX  
 AC AAB13639;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE C. trachomatis pmpG gene protein.  
 XX  
 KW Chlamydial infection; sexually transmitted disease;  
 KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
 KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;

KW coronary heart disease; antibacterial.  
 XX Chlamydia trachomatis.  
 OS WO200034483-A2.  
 XX 15-JUN-2000.  
 XX 08-DEC-1999; 99WO-US029012.  
 XX 08-DEC-1998; 98US-00208277.  
 PR 08-APR-1999; 99US-00288594.  
 PR 01-OCT-1999; 99US-00410568.  
 PR 22-OCT-1999; 99US-00426571.  
 XX (CORI-) CORIXA CORP.  
 XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
 XX WPI; 2000-431303/37.  
 DR Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
 PT comprises immunogenic portion of Chlamydia antigen, which comprises amino  
 PT acid sequence encoded by polynucleotide sequence.  
 XX Claim 2; Page 208-210; 256pp; English.  
 XX The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydiae are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention  
 XX Sequence 1006 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0 Length: 1006  
 Score: 5083.00 Matches: 978  
 Percent Similarity: 99.8% Conservative: 2  
 Best Local Similarity: 99.6% Mismatches: 0  
 Query Match: 64.5% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-701-844-1 (1-4435) x AAB13639 (1-1006)  
 QY 472 ATGGTCTCAAGAAATACGATGGGAGCGTTAACTGATCATCTTCCCTATCTGTT 531  
 DB 25 MetileProGlnGlyIleThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 44  
 QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTCGCGAGAGAGTTAACTTAAATGTT 591  
 DB 45 IleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64  
 QY 592 GACAAATCTATGCGCTTTGCTTAAAGTTGTTTGGGAACCTTATAGGAGTTTACT 651  
 DB 65 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84  
 QY 652 GTTTAGGAGAGACACTGTTGACCTTTTCGAGACATACGACTTCTCAAAATGGGCA 711  
 DB 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104  
 QY 712 GCTCTAGTAAATAGCGCTGCTGATGGAGTCTTTTACTATTGAGGCTTTTAAAGAAATATCC 771  
 DB 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124

QY 772 TTTTCCAATTGCAATTCATTACTTCCCGTACTGCTGCTGCAACGACTAATAAGGTAGC 831  
 DB 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144  
 QY 832 CAGACTCCGAGCAGCAACATCTACACCGCTCTAATGGTACTATTATTCTTAAACAGATCTT 891  
 DB 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleThrSerLysThrAspLeu 164  
 QY 892 TTGTTACTCAATTAATGAGAAGTTCTCATTTATAGTAATTTAGTCTCTGGAGATGGGGA 951  
 DB 165 LeuLeuLeuAsnAsnGlnLysPheSerPheThrSerAsnLeuValSerGlyAspGlyGly 184  
 QY 952 GCTATAGATGCTAAGAGCTTAAACGTTTCAAGGAATTTACAAAGCTTTGTCTTCCAGAA 1011  
 DB 185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 204  
 QY 1012 AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACAGGTTCTCTGCTATGGCT 1071  
 DB 205 AsnThrAlaGlnAlaAspGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224  
 QY 1072 AACGAGGCTCCTATTGCTTTGTAGCGAATGTTGAGGAGTAGTAAGAGGGGAGGATGCT 1131  
 DB 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 244  
 QY 1132 GCTGTTTCAAGGATGGCAGGAGTGTCTATCTTCAACAGAAAGATCCAGTAGTA 1191  
 DB 245 AlaValGlnAspGlyGlnGlyValSerSerSerThrThrGluAspProValVal 264  
 QY 1192 AGTTTTCAGAAATACCTCGGTAGATTTGATGGAGTAGTACCGGAGTAGGAGGAGG 1251  
 DB 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 284  
 QY 1252 ATTTTACTCTACGGGAACGCTCTTCTTCTGAATAATGGAACACCTTGTCTTCAACAAT 1311  
 DB 285 IleThrSerThrGlyAsnValAlaPheLeuAsnAsnGlyThrLeuPheLeuAsnAsn 304  
 QY 1312 GTTGCTTCTCTGTTTACATTTCTGCTAGCAACCAACCAAGTGGACAGCTTCTAATACG 1371  
 DB 305 ValAlaSerProValThrIleAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324  
 QY 1372 AGTAATAATTACGAGATGGAGAGCTATCTTCTGTAACAAATGGTGGCAGCAGATCC 1431  
 DB 325 SerAsnAsnThrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344  
 QY 1432 AATTAATCTGGATCAGTTTCTTGTGAGGAGGAGTAGTTTCTTATAGTCAATGTA 1491  
 DB 345 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 364  
 QY 1492 GCTGCTGGGAAAGGGGAGCTATTTATGCAAAAGCTCTCGGTGCTTAACCTGCGCCCT 1551  
 DB 365 AlaAlaGlyLysGlyAlaIleThrAlaLysLysLeuSerValAlaAsnCysGlyPro 384  
 QY 1552 GTACAAATTTTAAAGGAATATCGTAATGATGGTGGAGCGCATTTATTTAGGAGATCTGA 1611  
 DB 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleThrLeuGlyGluSerGly 404  
 QY 1612 GAGCTCAGTTTATCTGCTGATTTATGAGATATTATTTTCGATGGGAATCTTAAAGAAC 1671  
 DB 405 GluLeuSerLeuSerAlaAspThrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424  
 QY 1672 GCCAAAGAGAATGCTCCGATGTTAATGGCTAATCTGTCTCTCACAAAGCCATTTCCGATG 1731  
 DB 425 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444  
 QY 1732 GATCGGGAGGGGAAAATAACGATTAAGAGCTTAAAGAGGAGGATCATGATCTCTTAAAT 1791  
 DB 445 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464  
 QY 1792 GATCCCATCGAGATGCGCAACCGAATACCGCCGCGAGCTTCCAAACTTCTTAAAA 1851  
 DB 465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484

QY 1852 ATTAACGATGGTGAAGGATACACAGGGATATGTTTTGCTTAATGGAGCAGTACTTTG 1911  
Db 485 ILeAenAepGlyGluGlyThrGlyAepIleValPheAlaAenGlySerSerThrLeu 504  
QY 1912 TACCAAAATGTTACCATAGAGCAAGAGGATTGTTCTCGTGAAGAGCAAAATATCA 1971  
Db 505 TyrGlnAenValThrIleGluGlnIleValLeuArgGluValAlaAlaLeuSer 524  
QY 1972 GTGAATCTCTAAGTCAGACAGGTGGAGTCTGTATATGAAGCTGGGAGTACATGGAT 2031  
Db 535 ValAenSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAep 544  
QY 2032 TTTGTAACTCCAAACCACCAACAGCTCTCGCGCTTAATCAGTTGATCAGCTTTCC 2091  
Db 545 PheValThrProGlnProGlnProGlnProAlaAlaAenGlnLeuIleThrLeuSer 564  
QY 2092 AATCTGCATTTGCTCTCTCTCTCTTTGTTAGCAAAACAATGCAGTTACGAATCCTCTACC 2151  
Db 565 AenLeuHLeuSerLeuSerSerLeuLeuAlaAenAenAlaValThrAenProProThr 584  
QY 2152 AATCTCCAGCGCAAGATTCTCATCTGCAGTCATTGGTAGCACAACTGCTGTTCTGT 2211  
Db 585 AenProProAlaGlnAepSerHisProAlaValIleGlySerThrAlaGlySerVal 604  
QY 2212 ACAATTAGTGGCCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATGATGAT 2271  
Db 605 ThrIleSerGlyProIlePhePheGluAepLeuAepAepThrAlaTyrAepArgTyrAep 624  
QY 2272 TGGCTAGGTTCTAAATCAAAAATCAATGTCCTGAAATACAGTAGGAGCTAAAGCCCA 2331  
Db 645 TrpLeuGlySerAenGlyLysIleAenValLeuLysLeuGlnLeuGlyThrLysProPro 644  
QY 2332 GCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCCCTAAGTATGCTATCAAGA 2391  
Db 645 AlaAenAlaProSerAepLeuThrLeuGlyAenGluMetProLysTyrGlyTyrGlnGly 664  
QY 2392 AGCTGGAAGCTTGGTGGATCCTTAATACAGCAAAATATGCTCCTTATCTCTGAAGCT 2451  
Db 665 SerTrpLysLeuAlaTrpAepProAenThrAlaAenAenGlyProTyrThrLeuLysAla 684  
QY 2452 ACATCGACTAAACTGGGTATATCTGGCGCTGAGCGAGTAGTACTTCTTGGTCCAAAT 2511  
Db 685 ThrTrpThrLysThrGlyTyrAenProGlyProGluArgValAlaSerLeuValProAen 704  
QY 2512 AGTTATGGGGATCCATTTTAGATATACGATTCGCGCATTCAGCAATTCAGCAAGTGTG 2571  
Db 705 SerLeuTrpGlySerIleLeuAepIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724  
QY 2572 GATGGCGCTCTATTGTCGAGATTAAGGTTCTGGAGTTTCGAAATTTCTTCTATCAT 2631  
Db 725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAenPhePheTyrHis 744  
QY 2632 GACCGCATGCTTTAGTCAGGATATCGGTATATTAGTGGGGTATTCTCTTAGGAGCA 2691  
Db 745 AspaAsaPalaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764  
QY 2692 AACTCTACTTTGGATCATCGATGTTGCTAGCATTTACCAAGTATTTCGTAGATCT 2751  
Db 765 AenSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784  
QY 2752 AAAGATTATGATGTGCTGCTCCAAATCATCATGCTTGCATAGGATCCGTTTATCTATCT 2811  
Db 785 LysAepTyrValValCysArgSerAenHisAlaCysIleGlySerValTyrLeuSer 804  
QY 2812 ACCCAACAGCTTATGTCGATCCTATTGTCGGAGATGCGTTTATCCGTGCTAGCTAC 2871  
Db 805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAepAlaPheIleArgAlaSerTyr 824  
QY 2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTTCGAGAGGAGCGGATGTCGT 2931  
Db 825 GlyPheGlyAenGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAepValArg 844  
QY 2932 TGGGATATTAACGTCTGCTGGAGAGATTGGAGCGGGATTACCGATTGCTGATTACTCCA 2991

Db 845 TrpAenAenCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864  
QY 2992 TCTAAGCTCTATTGAATAGTTCGCTCTTTCTGTCGAAGCTGAGTTTCTTATGCCGAT 3051  
Db 865 SerLysLeuTyrLeuAenGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAep 884  
QY 3052 CATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGCAATTCAGAGCGGACATCTCCTA 3111  
Db 885 HisGluSerPheThrGluGluGlyAepGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904  
QY 3112 AATCTATCAGTTCCTGTTGGAGTGAAGTTTCATCGATGCTTCTAGTACACATCCCTAATAA 3171  
Db 905 AenLeuSerValProValGlyValLysPheAepArgCysSerSerThrHisProAenLys 924  
QY 3172 TATAGCTTTATGGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231  
Db 925 TyrSerPheMetAlaAlaTyrIleCysAepAlaTyrArgThrIleSerGlyThrGluThr 944  
QY 3232 AGCTCTCATCCCATCAAGACATGAGCAACAGATGCTTTCATTTAGCAAGACATGGA 3291  
Db 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaAargHisGly 964  
QY 3292 GTTGTGGTTAGAGGATCTATGATGCTTCTCTAAACAAGTAATATAAGAAGTATATGCCAT 3351  
Db 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAenIleGluValTyrGlyHis 984  
QY 3352 GGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGGAGTGCAGGAGTAGAGTC 3411  
Db 985 GlyArgTyrGluTyrArgAepAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004  
QY 3412 CGGTTTC 3417  
Db 1005 ArgPhe 1006  
RESULT 30  
ADD42756  
ID ADD42756 standard; protein; 670 AA.  
XX AC ADD42756;  
XX DT 15-JAN-2004 (first entry)  
XX DE Chlamydia pmpG passenger domain protein SEQ ID NO:169.  
XX KW Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;  
KW antinfertility; cardiant; antiarteriosclerotic; ophthalmological;  
KW vaccine; gene therapy; immune response; pelvic inflammatory disease;  
KW tubal obstruction; infertility; male infertility; ocular infection;  
KW blindness; acute respiratory tract infection; atherosclerosis;  
KW coronary heart disease.  
XX OS Chlamydia sp.  
XX PN WO2003041560-A2.  
XX PD 22-MAY-2003.  
XX PF 05-NOV-2002; 2002WO-US035624.  
XX PR 06-NOV-2001; 2001US-00012256.  
XX PR 05-DEC-2001; 2001US-00007693.  
XX PR 15-JUL-2002; 2002US-00197220.  
XX PA (CORI-) CORIXA CORP.  
XX PI Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;  
PI Probst P;  
XX WIPI; 2003-441771/41.  
XX DR N-PSDB; ADD42747.  
XX PT New DNA and proteins comprising a portion of a Chlamydia antigen, useful

PT for diagnosing or treating Chlamydial infections, particularly as  
 PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic  
 XX inflammatory disease.

XX Claim 2; SEQ ID NO 169; 275pp; English.

CC The present invention describes compounds and methods for diagnosing and  
 CC treating Chlamydial infection. Chlamydia polynucleotide and protein  
 CC sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,  
 CC antiarteriosclerotic and ophthalmological activities, and can be used in  
 CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,  
 CC compositions or methods from the present invention can be used for the  
 CC serodiagnosis or treatment of Chlamydial infections, particularly in  
 CC humans. The polynucleotides, proteins or compositions are particularly  
 CC useful for stimulating an immune response in a patient, or for  
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.  
 CC Specifically, the polynucleotides, proteins or compositions are useful as  
 CC vaccines for treating or preventing Chlamydial infections including  
 CC pelvic inflammatory disease (which results in tubal obstruction and  
 CC infertility in women), male infertility, ocular infection (which may  
 CC cause blindness), acute respiratory tract infections, atherosclerosis, or  
 CC coronary heart disease. The present sequence is used in the  
 CC exemplification of the present invention.

XX Sequence 670 AA;

Alignment Scores:

Pred. No.: 5,846-289 Length: 670  
 Score: 3336.50 Matches: 651  
 Percent Similarity: 98.8% Conservative: 11  
 Best Local Similarity: 97.2% Mismatches: 7  
 Query Match: 42.3% Indels: 1  
 DB: 7 Gaps: 1

US-10-701-844-1 (1-4435) x ADD42756 (1-670)

QY 463 CGAGAAATCATGTTCTCTCAAGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCC 522  
 DB 1 AlaGluIleMeIleProGInglyIleTyAspGlyGluThrLeuThrValSerPhePro 20  
 QY 523 TATACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTCTCGAGGAGATTAACTTA 582  
 DB 21 TyrThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeu 40  
 QY 583 AAAATCTTGACAAATCTATTGCGCTTGGCTTTAAGTTGTTGGGAACTATTAGGG 642  
 DB 41 LysAsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGly 60  
 QY 643 AGTTTACTGTTTATAGGAGAGACACTCGTTGACTTTTCGAGAACATACGACTTCTACA 702  
 DB 61 SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr 80  
 QY 703 AATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGAGTTTAA 762  
 DB 81 AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys 100  
 QY 763 GAATTATCTTTTCCAAATTCGAATTCATTATTCGCGTACTGCTGCTGCAACGACTAAT 822  
 DB 101 GluLeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaIleThrThrAsn 120  
 QY 823 AAGGCTAGCCAGACTCCGAGCAACATCTACACCGCTTAATCGTACTATTATTTCTAAA 882  
 DB 121 AsnGlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLys 140  
 QY 883 ACAGATCTTTTGTACTCAATAATAGAACTTCTCATTTCTATAGTAATTTAGTCTCTGGA 942  
 DB 141 ThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTySerAsnLeuValSerGly 160  
 QY 943 GATGGGGAGCTATAGATGCTTAAGCTTAAACGGTTCAAGGAATTAGAGCTTTGTGTC 1002  
 DB 161 AspGlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysVal 180  
 QY 1003 TTCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAGTAGTCACCAGTTTCTCT 1062

181 PheGlnGluAsnThrAlaGlnAlaAspGlyAlaCysGlnValValThrSerPheSer 200  
 1063 GCTATGGCTACAGAGCTCTATTGCTTTCTAGCGAATGTTGCGAGGAGTAGAGGGGA 1122  
 201 AlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsnValAlaGlyValArgGlyGly 220  
 1123 GGGATTGCTGTTCTAGGATGGGAGGAGTGTCTCATCTTCTTCAACAGAGAT 1182  
 221 GlyIleAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAsp 240  
 1183 CCAGTAGTAAGTTTTCAGAAATCTCGGTAGAGTTTGTGGAAGTACGAGCCGAGTA 1242  
 241 ProValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgVal 260  
 1243 GGAGGAGGATTTACTCTCTAGCGAACTGCTTCTCTGAATAATGGAATAACCTTGT 1302  
 261 GlyGlyGlyIleTySerTyGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPhe 280  
 1303 CTCACAAATGTTGCTTCTCTGTTTACATTGCTGTAAACCAACCAAGTCGACAGGT 1362  
 281 LeuAsnAsnValAlaSerProValTyIleAlaAlaGluGlnProThrAsnGlyGlnAla 300  
 1363 TCTAATACGAGTAATAATTACGAGATCGAGAGCTATCTTCTGTAAAGATGTCGCA 1422  
 301 SerAsnThrSerAspAsnTyArgAspGlyGlyAlaIlePheCysLysAsnGlyAlaGln 320  
 1423 ---GCAGGATCCAAATACTCTGATCAGTTTCTTGTGATGGAGGGAGTAGTTTCTTT 1479  
 321 AlaAlaGlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePhe 340  
 1480 AGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTATTGCCAAAAGTCTCGTGTGT 1539  
 341 SerSerAsnValAlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAla 360  
 1540 AACTGTGCGCTGTACAATTTTAAAGGAATATCGCTAATATGATGGTGGAGCGATTATTTA 1599  
 361 AsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsnAspGlyGlyAlaIleTyLeu 380  
 1600 GGAGATCTGGAGAGCTCAGTTTATCTGCTGATATGAGATATTATTTTCCATGGGAAT 1659  
 381 GlyGluSerGlyGluLeuSerLeuSerAlaAspTyArgLysIleIlePheAspGlyAsn 400  
 1660 CTTAAAAGAACCAACCAAGAGATGCTCCGATGTTTAAATGGCGTAATCTGTCTCTCACAA 1719  
 401 LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln 420  
 1720 GCCATTTCCATGGGATCGGAGGAGGAAATAACGACATTAAGAGCTAAAGCAGGCGATCAG 1779  
 421 AlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGln 440  
 1780 ATTCTCTTAATGATCCCATCGAGATGGCAACCGGAAATAACCGCCAGCGAGTCTCC 1839  
 441 IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer 460  
 1840 AAATCTCTAAAAATTAACGATGCTGAAGGATACACAGGCGGATATTGTTTTGCTAATGGA 1899  
 461 GluProLeuLysIleAsnAspGlyGluGlyTyThrGlyAspIleValPheAlaAsnGly 480  
 1900 AGCAGTACTTTGTACCAAAATGTTACGATAGAGCAAGGAGGATTTCTTCTGCGAAAAG 1959  
 481 AsnSerThrLeuTyGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLys 500  
 1960 GCAAAATTAATCAGTGAATTTCTTAAGTCAGACAGGTGGAGCTGTATATGAGAGCTGG 2019  
 501 AlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyMetGluAlaGly 520  
 2020 AGTACATGGGATTTGTAACTTCACTCCACAAACCAACCAACAGCTCTCTCCGCTAATCAGTTG 2079  
 521 SerThrLeuAspPheValThrProGlnProProGlnGlnProProAlaAlaAsnGlnLeu 540  
 2080 ATCACGCTTTTCCAATCTGCAATTTGCTCTTCTTCTTTTGTAGCAAAACAAATGAGTTACG 2139

Db 541 IleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThr 560  
QY 2140 AATCCTCTTACCAATCTCTCAGCGCAGAGATTCTCATCTCTGAGTCATTGGTAGCACAACT 2199  
Db 561 AsnProProThrAsnProProAlaGlnAspSerHisProAlaIleIleGlySerThrThr 580  
QY 2200 GCTGGTCTCTGTACAAATAGTAGGCGCTATCTTTTGGAGGATTTGGATGATACAGCTTAT 2259  
Db 581 AlaGlySerValThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyr 600  
QY 2260 GATAGGTATGATGGCTAGGTTCTTAATCAAAAAATCAATGCTCCTGAAATTCACGTTAGGG 2319  
Db 601 AspArgTyrAspTyrLeuGlySerAsnGlnIleAspValLeuLeuLysLeuGlnLeuGly 620  
QY 2320 ACTAAGCCCCCAGCTAATGCCCCCATCAGATTGACTCTAGGGAATGAGATGCCCTAAGTAT 2379  
Db 621 ThrGlnProSerAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyr 640  
QY 2380 GCGTATCAAGGAAGCTGGAAGCTTGGCGTGGATCTCTAATACAGCAAAATAATGGTCCTTAT 2439  
Db 641 GlyTyrGlnGlySerTyrLeuAlaTyrAspProAsnThrAlaAsnAsnGlyProTyr 660  
QY 2440 ACTGTAAAGCTACATGCACTAAAAGCTGGG 2469  
Db 661 ThrLeuLysAlaThrTyrThrLysThrGly 670

RESULT 31

ID AAY37238 standard; protein; 524 AA.

XX AAY37238;

XX 07-OCT-1999 (first entry)

XX Chlamydia trachomatis cellular envelope protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;  
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX Chlamydia trachomatis.

XX WO9928475-A2.

XX 10-JUN-1999.

XX 27-NOV-1998; 98WO-IB001939.

XX 28-NOV-1997; 97FR-00015041.

XX 17-DEC-1997; 97FR-00016034.

XX 04-NOV-1998; 98US-0107077P.

XX (GEST ) GENSET.

XX Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis.

XX Disclosure; Page 996-997; 1755pp; English.

XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
CC of Chlamydia trachomatis (see AAY01425). The polypeptides can be used as  
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
CC can also be used to control growth of the microorganism. Chlamydia  
CC trachomatis is responsible for a large number of diseases, e.g. eye  
CC diseases such as conventional trachoma, nonendemic trachoma,  
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis, and  
CC perihhepatitis, bartholinitis; pneumopathy in breast feeding infants; and  
CC venereal lymphogranulomatosis. The polypeptides of the invention may be

CC of use in treating these diseases  
XX SQ Sequence 524 AA;  
Alignment Scores: 1.3e-223 Length: 524  
Pred. No.: 2606.00 Matches: 518  
Score: 99.2% Conservative: 0  
Percent Similarity: 99.2% Mismatches: 4  
Best Local Similarity: 33.1% Indels: 1  
Query Match: 2 Gaps: 0  
DB: 2

US-10-701-844-1 (1-4435) x AAY37238 (1-524)

QY 625 TTTGGGAACCTATTAGGGAGGTTTACTGTGTTTATGGGAGAGGACACTCGTTGACTTTTCGAG 684  
Db 3 PheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPheGlu 22  
QY 685 AACATACGAGCTTCTCAAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTT 744  
Db 23 AsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPhe 42  
QY 745 ACTATTGAGGGTTTAAAGAAATTATCTTTTCCAAATTCGAATTCATTACTTTCGCTACTG 804  
Db 43 ThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeu 62  
QY 805 CCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACACCGTCTAAT 864  
Db 63 ProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrSerThrProSerAsn 82  
QY 865 GGTACTATTATCTTAAACAGATCTTTTGTACTCAATATAGAGATTTCTCATTTCTAT 924  
Db 83 GlyThrIleTyrSerLysThrAspLeuLeuLeuAsnGluLysPheSerPheTyr 102  
QY 925 AGTAATTAGCTCTCGGAGATGGGGAGCTATAGATCTAAGAGCTTAACGGTTCACAGGA 984  
Db 103 SerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGlnGly 122  
QY 985 ATTAGCAAGCTTTGTGCTTCCAAAGAAATACTCTCAAGCTGATGGGGAGCTTGTCAA 1044  
Db 123 IleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGln 142  
QY 1045 GTAGTCACAGTTTCTGCTATGCTTAACGAGGCTCTATTGCTTTGTAGCGAATGTT 1104  
Db 143 ValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsnVal 162  
QY 1105 GCAGGAGTAAGAGGGGGAGGATTCGCTGCTCAGGATGGCGAGCGAGGTGTCATCA 1164  
Db 163 AlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSerSer 182  
QY 1165 TCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTTCCAGAAATACCTCGGTAGAGTTGAT 1224  
Db 183 SerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPheAsp 202  
QY 1225 GGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTTACGGGAACGTTGCTTTCTCTGAT 1284  
Db 203 GlyAsnValAlaArgValGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsn 222  
QY 1285 AATGGAAAAACCTTGTGTTTCTCAACAAATGTTGCTTCTCTGTTTACATTGCTGCTAAGCAA 1344  
Db 223 AsnGlyLysThrLeuPheLeuAsnAsnValAlaPheProValTyrIleAlaAlaLysGln 242  
QY 1345 CCAACCAAGTGACAGGCTTCTAATACGAGTAGTAATTAATTCGGAGATGGAGGAGCTATCTTC 1404  
Db 243 ProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIlePhe 262  
QY 1405 TGTAAGAAATGTCGCAAGACGAGATCCAAATACCTCGGATCAGTTCTTCCTGATGAGAG 1464  
Db 263 CysLysAsnGlyAlaGlnValGlySerAsnAsnSerGlySerValSerPheAspGlyGlu 282  
QY 1465 GGAGTAGTTTCTTTTATAGGCAATCTAGCTGCTGGGAAAGGGGAGCTATTTATGCAAAA 1524  
Db 283 GlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLys 302





```
Db 191 GlnGluAenThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSerAla 200
Qy 1066 ATGGCTAACGAGGCTCTATTGCTTCTAGCAATGTTGCAGAGTAGTAAAGGGGAGGG 1125
Db 201 MetAlaAenGluAlaProIleAlaPheValAlaAenValAlaGlyValArgGlyGlyGly 220
Qy 1126 ATTGCTGCTGTTACGAGTGGGAGGAGGAGTGTATCATCTTCAACAGAGATCCA 1185
Db 221 IleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrGluAspPro 240
Qy 1196 GTAGTAAGTTTCCAGAACTACGCGCTAGGTTGATGGAGCTAGCCGAGTAGGA 1245
Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
Qy 1246 GGAGGATTTACTCTACGGGAACCTGCTTCTCGAATAATGGAATAACCTTGTCTC 1305
Db 261 GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAenAsnGlyLysThrLeuPheLeu 280
Qy 1306 AACATGTTGCTTCTCTGTTTACATTCGCTGAAGCAACCAAGTGGACAGCTTCT 1365
Db 281 AenAenValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
Qy 1366 AATACGACTAATAATACGAGATGAGGAGCTATCTCTGTAAGATGTTGGCGAAGCA 1425
Db 301 AenThrSerAenAsnTyrGlyAspGlyGlyAlaIlePheCysLysAenGlyValAlaGlnAla 320
Qy 1426 GGATCCAATAACTCTGGAGTCAGTTCTCTTGTGAGGAGGAGTAGTATTTCTTAGTAGC 1485
Db 321 GlySerAenAenSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340
Qy 1486 AATGTAGCTGTGGAAAGGGGAGCTATTATATGCCAAAGCTCTCGTTGCTAACTGT 1545
Db 341 AenValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAenCys 360
Qy 1546 GGCCTGTACATTTTAAAGATATCGCTAATGATGTTGGGCGCATTTATTAGAGAA 1605
Db 361 GlyProValGlnPheLeuArgAsnIleAlaAenAspGlyGlyAlaIleTyrLeuGlyGlu 380
Qy 1606 TCTGAGAGCTCAGTTTATCTCTGATATGAGATATTATTTTCGATGGGAATCTTAAA 1665
Db 381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400
Qy 1666 AGAACAGCCAAAGAGATGCTCCGATGTTAATGGCGTAACGTGTCTCTCAAGCCATT 1725
Db 401 ArgThrAlaLysGluAenAlaAlaAspValAenGlyValThrValSerSerGlnAlaIle 420
Qy 1726 TCGATGGGATCGGAGGGGAAATAACGACATTAAAGCTTAAGCAGGCGGATCAGATTCTC 1785
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
Qy 1786 TTTAATGATCCCATCGAGATGGCAAGCAAGCAATACACCGCAGCGAGTCTTCCAACTT 1845
Db 441 PheAenAspProIleGluMetAlaAenGlyAenAenGlnProAlaGlnSerSerLysLeu 460
Qy 1846 CTPAAAATTAAAGATGGTGAAGATACACAGGGATATTGTTTTGCTTAATGGAAGCAGT 1905
Db 461 LeuLysIleAenAspGlyGluGlyTyrThrGlyAspIleValPheAlaAenGlySerSer 480
Qy 1906 ACTTTGTACCAAAATGTTACGATAGCAGGAGGAGGATTTCTTCGTGAAGAAGCCAAA 1965
Db 481 ThrLeuTyrGlnAenValThrIleGluGlnGlyArgIleValLeuArgGlnLysAlaLys 500
Qy 1966 TTATCAGTGAATTCT 1980
Db 501 LeuSerValAspSer 505
```

RESULT 33

ID AAY16751

XX AAY16751 standard; protein; 158 AA.

AC AAY16751;

XX

```
DT 21-JUL-1999 (first entry)
XX Chlamydia HMW protein fragment.
DE
XX Chlamydia; high molecular weight protein; HMW protein; urethritis;
KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
KW salpingitis; tubal occlusion; infertility; cervical cancer;
KW arteriosclerosis; atherosclerosis.
XX Chlamydia sp.
OS
XX W09917741-A1.
PN
XX 15-APR-1999.
PD
XX 01-OCT-1998; 98WO-US020737.
PF
XX 02-OCT-1997; 97US-00942596.
PR
XX (ANTE-) ANTEX BIOLOGICS INC.
PA
XX Jackson JW, Pace JL;
PI
XX WPI; 1999-287659/24.
DR
XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
PT cervical cancer.
PT Claim 5; Page 133-135; 141pp; English.
PS
XX The invention relates to an isolated Chlamydia species high molecular
CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
CC be used for preventing, treating or ameliorating a disorder related to
CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
CC cancer, infertility, arteriosclerosis and atherosclerosis. The products
CC can also be used for detection and diagnosis. Sequences AAY16740-Y16752
CC represent Chlamydia HMW protein fragments
XX Sequence 458 AA;
SQ
Alignment Scores:
Pred. No.: 1,01e-200 Length: 458
Score: 2350.00 Matches: 458
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.8% Indels: 0
DB: 2 Gaps: 0
US-10-701-844-1 (1-4435) x AAY16751 (1-458)
Qy 1030 GGGGAGCTGTCAAGTAGTCACCAATTTCTCTGCTATGGCTAACGAGCTCTATTGCC 1089
Db 1 GlyGlyAlaCysGlnValThrSerPheSerAlaMetAlaAenGluAlaProIleAla 20
Qy 1090 TTTGTAGCGCAATGTTGCAGAGTAGTAAGAGGGAGGAGATTGCTGCTGTTCAGGATGGCAG 1149
Db 21 PheValAlaAenValAlaGlyValArgGlyGlyGlyIleAlaValGlnAspGlyGln 40
Qy 1150 CAGGAGGTGTCAATCATCTACTTCAACAGAAAGTCCAGTAGTAGTATTTTCCAGAAATACT 1209
Db 41 GlnGlyValSerSerSerThrThrGluAspProValValSerPheSerArgAsnThr 60
Qy 1210 GCGGTAGAGTTTGATGGGAACGTAGCCGCTAGGAGGAGGAGGATTTACTCTACCGGAAC 1269
Db 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAen 80
Qy 1270 GTTGTCTTCTCGAATAATGGAATAACCTTGTCTTCTCAACAATGTTGCTTCTCTGTTTAC 1329
Db 81 ValAlaPheLeuAenAsnGlyLysThrLeuPheLeuAenAsnValAlaSerProValTyr 100
```

QY 1330 ATTGCTGCTAGCAACCAAGTGGACAGGCTTCTAATACAGTAATAATTACGAGAT 1389  
 DB 101 IleaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAsp 120  
 QY 1390 GGAGAGCTATCTCTGTAAGAATGCTGGCAAGCAGGATCAATAACTCTCGATCAGTT 1449  
 DB 121 GlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal 140  
 QY 1450 TCCCTTGTAGAGAGAGAGTACTTTCTTTAGTACCAATGTAAGTCTGGGAAAGGGGA 1509  
 DB 141 SerPheAspGlyGluGlyValPhePheSerSerAsnValAlaAlaGlyGlyGly 160  
 QY 1510 GCTATTATGCCAAAAGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1569  
 DB 161 AlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn 180  
 QY 1570 ATCGCTAAATGATGCTGGAGCGATTTATTTAGGAGATCTGGAGAGCTCAATTTATCTGCT 1629  
 DB 181 IleaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAla 200  
 QY 1630 GATTATGGAGATATTTTTCGATGGGAATCTTTAAAGAACAGCCAAAGAGATGCTGCC 1689  
 DB 201 AspTyrGlyAspIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAla 220  
 QY 1690 GATGTTAATGGCTAATGCTGCTCCTCAAGCCATTTTCATGCGATCGGAGCGGAATAA 1749  
 DB 221 AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIle 240  
 QY 1750 ACACATTAAGAGTAAAGCAGGCGATCTCTTTAATGATCCATCGAGATGGCA 1809  
 DB 241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAla 260  
 QY 1810 AACGGAATAACAGCAGCGAGTCTTCCAACTTTCTAAATTAACCATGCTGAAGGA 1869  
 DB 261 AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLysIleAsnAspGlyGluGly 280  
 QY 1870 TACACAGGGATATGTTTTGCTAATGAGCAGTACTTTGTACCAAAATGTCAGATA 1929  
 DB 281 TyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIle 300  
 QY 1930 GAGCAGGAGGATGTTCTCTCTGAAAGGCAAAATATCATGTAATCTCTAAGTCAG 1989  
 DB 301 GluGlnGlyArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGln 320  
 QY 1990 ACAGGTGGAGTCTGTATATGAGAGCTGGAGTACATGGATTTTGTAACTCCACACCA 2049  
 DB 321 ThrGlyGlySerLeuTyrMetGluAlaGlySerThrTrpAspPheValThrProGlnPro 340  
 QY 2050 CCACACAGCTCTCTGCGCTAATCATGTTGATCAGCTTTCCAACTCTGCATTTCTCTT 2109  
 DB 341 ProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu 360  
 QY 2110 TCTTCTTTGTAGCAACATGAGTTTACCAATCTCTACCAATCTCTCAGCGCAGAT 2169  
 DB 361 SerSerLeuLeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAsp 380  
 QY 2170 TCTCATCTCTGAGTATGTTAGCACAACCTGCTGTTCTGTATCAATTTAGTGGCCTATC 2229  
 DB 381 SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIle 400  
 QY 2230 TTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTAATCAA 2289  
 DB 401 PhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGln 420  
 QY 2290 AAAATCAATGCTCTGAAATACAGTATAGGACTAAGCCCAAGCTAATGCCCATCAGAT 2349  
 DB 421 LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAsp 440  
 QY 2350 TTGACTCTAGGAAATGAGATGCTTAACTATGCTGCTATCAAGGAAGCTTGAAGCTT 2403  
 DB 441 LeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeu 458

RESULT 34

AAY16752  
 ID AAY16752 standard; protein; 325 AA.  
 AC AAY16752;  
 XX 21-JUL-1999 (first entry)  
 XX Chlamydia HMW protein fragment.  
 XX Chlamydia; high molecular weight protein; HMW protein; urethritis;  
 KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;  
 KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;  
 KW salpingitis; tubal occlusion; infertility; cervical cancer;  
 KW arteriosclerosis; atherosclerosis.  
 XX Chlamydia sp.  
 XX WO9917741-Al.  
 XX 15-APR-1999.  
 XX 01-OCT-1998; 98WO-US020737.  
 XX 02-OCT-1997; 97US-00942596.  
 XX (ANTE-) ANTEX BIOLOGICS INC.  
 XX Jackson JW, Pace JL;  
 XX WPI; 1999-287659/24.  
 XX New Chlamydia protein useful for treating conjunctivitis, urethritis and  
 PT cervical cancer.  
 XX Claim 5; Page 135-136; 14lpp; English.  
 CC The invention relates to an isolated Chlamydia species high molecular  
 CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as  
 CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can  
 CC be used for preventing, treating or ameliorating a disorder related to  
 CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,  
 CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,  
 CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical  
 CC cancer, infertility, arteriosclerosis and atherosclerosis. The products  
 CC can also be used for detection and diagnosis. Sequences AAY16740-Y16752  
 CC represent Chlamydia HMW protein fragments  
 XX Sequence 325 AA;  
 Alignment Scores:  
 Pred. No.: 9,67e-146 Length: 325  
 Score: 1735.00 Matches: 325  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 22.0% Indels: 0  
 DB: Gaps: 0  
 US-10-701-844-1 (1-4435) x AAY16752 (1-325)  
 QY 2443 CTGAAAGCTACATGAGTAAACTGGGTATTAATCTGGGCTGAGCGAGTAGCTTCTTTG 2502  
 DB 1 LeuLysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeu 20  
 QY 2503 GTTCCAAATAGTTTATGGGATCCATTTAGATATAGATCTGGCATTCAGCAATTCAA 2562  
 DB 21 ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln 40  
 QY 2563 GCAAGTGTGATGGCGCTTATCTCGAGGATTTAGGTTCTGAGTTTCGAATTTTC 2622  
 DB 41 AlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPro 60  
 QY 2623 TTCTATCATCACCGCGATGCTTTAGGTCAGGATATCGGTATATAGTGGGGTATTATTC 2682



Db 141 llePheTrpAspAsnSerValGlyTyrSerProLeuSerThrValProThrSerSer 160  
Qy 3922 ACTCCGCTGCTCCCAACAGTATGATGCTCGGAAGGCTCTATTTTCTGTAGAGACT 3981  
Db 161 ThrProAlaProThrValSerAspAlaArgLysGlySerIlePheSerValGluThr 180  
Qy 3982 AGTTTGGAGATCTCAGGCTCAAAAAGGGGTCATGTTTCGATAAATCCCGGGAATTC 4041  
Db 181 SerLeuGluIleSerGlyValLysGlyValMetPheAspAsnAlaGlyAsnPhe 200  
Qy 4042 GGAACAGATTTTTCGAGGTAAAGATAATAATATGCTGGTGGAGGCGAGTCGGTCCGC 4101  
Db 201 GlyThrValPheArgGlyLysAsnAsnAlaGlyGlySerGlySerAl 220  
Qy 4102 TACACATCAAGTACGACTTTTACGTTAAACCTGTAAGGAAAGTTTCTTTCACAGA 4161  
Db 220 aThrProSerSerThrThrPheThrValLysAsnCysLysGlyValSerPheThrAs 240  
Qy 4162 TAACTAGCTCTTCGAGGCGGAGTGGTTTATAAAGCATTTGCTTTTCAAGACAA 4221  
Db 240 pAsnValAlaSerCysGlyGlyValValLysGlyValLysGlyValLysPheLysAspAs 260  
Qy 4222 TGAAGGAGCATATTTCTCCGAGGGAACACAGCATACGATGATTTTAAGGATTTCTGCTC 4281  
Db 260 nGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspLeuArgIleLeuAlaAl 280  
Qy 4282 TACTAATCAGGATCAGAATACGAGACAGAGCGGTGAGGAGTATTTGCTCTCCAGA 4341  
Db 280 aThrAsnGlnAepGlnAsnThrGluThrGlyGlyGlyGlyValLysCysSerProAs 300  
Qy 4342 TGAATCTGAAGTTTGAAGGCAATAAAGTTCTATTGTTTGTATTACAACTTTCCAAA 4401  
Db 300 pAepSerValLysPheGluGlyAsnLysGlySerIleValPheAspTyrAsnPheAlaL 320  
Qy 4402 AGGCAGGCGGGAAGCATCTCAACGAAAGAAATTC 4435  
Db 320 sGlyArgGlyGlySerIleLeuThrLysGluPhe 331  
RESULT 36  
AAG83274  
ID AAG83274 standard; protein; 631 AA.  
XX AC AAG83274;  
DT 05-SEP-2001 (first entry)  
XX DE Chlamydia trachomatis PmpH(N-term) fusion protein.  
XX Chlamydia; vaccine; infection; fusion protein; antigen;  
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;  
KW acute respiratory tract infection; Capi; CT529; OMCB;  
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.  
XX OS Chlamydia trachomatis.  
XX WO200140474-A2.  
XX 07-JUN-2001.  
XX 04-DEC-2000; 2000WO-US032919.  
XX 03-DEC-1999; 99US-00454684.  
XX 19-APR-2000; 2000US-00556877.  
XX 20-JUN-2000; 2000US-00598419.  
XX (CORI-) CORIXA CORP.  
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;  
XX WPI; 2001-374831/39.  
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic  
XX inflammatory disease, trachoma, acute respiratory tract infections,  
PT

PT atherosclerosis and heart disease.  
XX Claim 70; Page 274-276; 295pp; English.  
XX The present sequence is provided in a specification relating to compounds  
CC and methods for the treatment and diagnosis of chlamydial infection. The  
CC compounds provided include polypeptides and fusion proteins comprising  
CC immunogenic portions of Chlamydia antigens and DNA sequences encoding  
CC such polypeptides. They are useful for vaccinating against chlamydial  
CC infection, which causes pelvic inflammatory disease, trachoma, acute  
CC respiratory tract infections, atherosclerosis and heart disease  
XX Sequence 631 AA;  
SQ  
Alignment Scores:  
Pred. No.: 1,21e-132 Length: 631  
Score: 1590.50 Matches: 316  
Percent Similarity: 88.7% Conservative: 7  
Best Local Similarity: 86.8% Mismatches: 19  
Query Match: 20.2% Indels: 22  
DB: 4 Gaps: 4  
US-10-701-844-1 (1-4435) x AAG83274 (1-631)  
Qy 3391 TTGAGTGCAGGAAGTAGAGTCGCGTTCTAAATAATAT-----TGGTTAGATAGTTAA 3441  
Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisTrpArgPro----- 146  
Qy 3442 GTGTAGCGATGCTTTTCTTTGAGATCTACATCATTTTGTGTTTGTAGCTTGTGTTGT 3501  
Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164  
Qy 3502 TCCTATTTCGTATGATTCGCGAGCTCTCTCAAGTGTTAACGCTAATTAACCACTCTCT 3561  
Db 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrPro 184  
Qy 3562 TTTAAGGAGACGATGTTTACTTGAATGGAGACTGCGCTTTTGTCAATGTCTATGAGGA 3621  
Db 185 PheLysGlyAspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGly 204  
Qy 3622 GCTGAAGAAGGTTCTGATTTATCTCAGCTAATGGGACAAATTAACGATTACCGACAAAC 3681  
Db 205 AlaGluAsnGlySerIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 224  
Qy 3682 CATACATTATCATTTTACAGATTCTCAAGGCCAGTTCTTCAAAATTTATGCTTTCATTTC 3741  
Db 225 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSer 244  
Qy 3742 GCAGGAGACACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCT 3801  
Db 245 AlaGlyGluThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSer 264  
Qy 3802 TCGGAGAAAGGAATGATCTCCGGGAAACCGTCAGTATTTCCGAGGAGGCGGAGTG 3861  
Db 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284  
Qy 3862 ATTTTCTGGGATAACTCCGTTGGGTATTTCTCTTTTATCTACTGTCGCAACTCATCATCA 3921  
Db 285 IlePheTrpAspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304  
Qy 3922 ACTCCGCTGCT-----CCACAGATTAGT 3945  
Db 305 ThrProProAlaProAlaProAlaProAlaAlaSerSerSerLeuSerProThrValSer 324  
Qy 3946 GATGCTCGGAAACGGTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAA 4005  
Db 325 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 344  
Qy 4006 AAAGGGTCTATGTTCCGATATATATGCGGAAATTCGGAACAGTTTTTTCGAGGTAAAGAT 4065  
Db 345 LysGlyValMetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSer 364  
Qy 4066 AATAAATATGCTGTTGGTGGAGGCGAGTGGGTTCGCTTACACCATCAAGTACGACTTTTAC 4125

Db 365 AenAenAenAlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheTh 382  
 QY 4126 AGTTAAACCTGTAAAGGAAAGTTCTTTTCACAGATAACGTAGCCTCTTCGGAGGCGG 4185  
 Db 382 rValLysAenCysLysGlyLysValSerPheThrAaspAenValAlaSerCysGlyGlyG1 402  
 QY 4186 AGTGGTTTAAAGGCATTTGCTTTTCAAGACATGAAGAGGACATATCTTCGGAGG 4245  
 Db 402 yValValLysGlyThrValLeuPheLysAaspAenGluGlyGlyLysPhePheArgG1 422  
 QY 4246 GAACACAGCATACGATGATTTAAGCATTTCTGCTCTACTAATCAGGATCAGAATACGGA 4305  
 Db 422 yAenThrAlaLysAaspAenLysGlyLysLeuAlaAlaThrSerArgAaspGlnAenThrG1 442  
 QY 4306 GACAGAGCGGTGGAGGAGTTATTTGCTCTCCAGATGATTTCTGAAAGTTTGAAGGCAA 4365  
 Db 442 uThrGlyGlyGlyGlyValLysCysSerProAaspAaspSerValLysPheGluGlyA8 462  
 QY 4366 TAAAGGTTCTATGTTTGTGATTACAACTTTGCAAAAGGACAGGCGGAGGATCCTAAC 4425  
 Db 462 nLysGlySerIleValPheAaspTyrAaspPheAlaLysGlyArgGlyGlySerIleLeuTh 482  
 QY 4426 GAAAGCAATTC 4435  
 Db 482 rLysGluPhe 485

RESULT 37  
 ABB94245  
 ID ABB94245 standard; protein; 631 AA.  
 XX ABB94245;  
 AC ABB94245;  
 DT 05-JUN-2002 (first entry)  
 XX Chlamydia trachomatis protein sequence SEQ ID NO:325.  
 DB Chlamydia infection; Chlamydia; vaccine; detection; diagnosis; antigen;  
 KW antibacterial; immunostimulant; immune response;  
 KW Chlamydia-specific T-cell response.  
 XX Chlamydia trachomatis.  
 XX WO200208267-A2.  
 PN 31-JAN-2002.  
 XX 20-JUL-2001; 2001WO-US023121.  
 XX 20-JUL-2000; 2000US-00620412.  
 PR 23-APR-2001; 2001US-00841132.  
 XX (CORI-) CORIXA CORP.  
 PA Fling SP, Skeiky YAW, Probst P, Bhatia A;  
 PI WPI; 2002-179901/23.  
 XX Novel compositions comprising Chlamydia Cap1 protein and its use in the  
 PT treatment of Chlamydia infection.  
 XX Example 10; Page 300-302; 537pp; English.  
 CC The present invention describes compositions comprising a Chlamydia Cap1  
 CC protein and methods for the diagnosis and therapy of Chlamydia infection.  
 CC Chlamydia DNA and protein sequences from the present invention can have  
 CC antibacterial and immunostimulant activities, and can be used in  
 CC vaccines. Compounds from the present invention can be used for eliciting  
 CC an immune response, specifically stimulating a Chlamydia-specific T-cell  
 CC response or inhibiting the development of a Chlamydia infection in an  
 CC animal. Methods from the present invention can be used for detecting the  
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells  
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia

CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent  
 CC sequences used in the exemplification of the present invention  
 XX SQ Sequence 631 AA;  
 Alignment Scores:  
 Pred. No.: 1,21e-132 Length: 631  
 Score: 1590.50 Matches: 316  
 Percent Similarity: 88.7% Conservative: 7  
 Best Local Similarity: 86.8% Mismatches: 19  
 Query Match: 20.2% Indels: 22  
 DB: Gaps: 4  
 US-10-701-844-1 (1-4435) x ABB94245 (1-631)  
 QY 3391 TTGAGTCGAGAGTAGAGTCGGCTTCTAAATAATAT-----TGTTAGATAGTTAA 3441  
 Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisTipArgPro----- 146  
 QY 3442 GTGTTAGCGATCGCTTTTCTTTGAGATCTACATCATTTCTTTTGTGTTGTGTGTGT 3501  
 Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164  
 QY 3502 TCCTATTCTGATGATTCGCGAGCTCTCTCAAGTGTTAAGCCCTAAATGTAAACCACTCTCT 3561  
 Db 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAenValThrThrPro 184  
 QY 3562 TTTAAGGAGACGATGTTTACTTGAATGAGACATGCGCTTTTGTCAATGTCTATCAGGA 3621  
 Db 185 PheLysGlyAaspValTyrLeuAenGlyAaspCysAlaPheValAenValTyrAlaGly 204  
 QY 3622 GCTGAAGAGGTTGATTTATCTCAGCTAATCGGACAAATTTAAGATTTACCGGACAAAAC 3681  
 Db 205 AlaGluAenGlySerIleSerAlaAenGlyAaspAenLeuThrIleThrGlyGlnAen 224  
 QY 3682 CATACATTATCATTTACAGATTTCTCAAGGCGCAGTTCTTCAAAATATATGCTCTCATTTCA 3741  
 Db 225 HisThrLeuSerPheThrAaspSerGlnGlyProValLeuGlnAenTyrAlaPheIleSer 244  
 QY 3742 GAGGAGAGACATTTACTCTGAGAGATTTTTCGAGCTGTGATGTTCTCGAAAATGTTTCT 3801  
 Db 245 AlaGlyGluThrLeuThrLeuLysAaspPheSerSerLeuMetPheSerLysAenValSer 264  
 QY 3802 TCGCGAGAAAGGGAATGATCTCCGGGAAACCGTGAATTTCCGGAGGAGGCGCAAGTG 3861  
 Db 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284  
 QY 3862 ATTTTCTGGGATACTCCGTGGGTATTTCTCTTTATCTTACTGTGCCAACCTCATCATCA 3921  
 Db 285 IlePheThrAaspAenSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304  
 QY 3922 ACTCGCCTGCT-----CCACAGTTAGT 3945  
 Db 305 ThrProProAlaProAlaProAlaProAlaSerSerSerLeuSerProThrValSer 324  
 QY 3946 GATGCTCGGAAAGGCTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCTCAA 4005  
 Db 325 AspAlaLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 344  
 QY 4006 AAAGGGGTGATGTCGATAATAATCCGGGAATTTTCGGAACAGTTTTCAGGTGAAGAT 4065  
 Db 345 LysGlyValMetPheAaspAenAlaGlyAenPheGlyThrValPheArgGlyAenSer 364  
 QY 4066 AATAATAATGCTGGTGGAGGAGTGGTTCCTACCATCAATCAAGTACGACTTTTAC 4125  
 Db 365 AenAenAenAlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheTh 382  
 QY 4126 AGTTAAACCTGTAAAGGAAAGTTCTTTTCACAGATAACGTAGCCTCTTCGGAGGCGG 4185  
 Db 382 rValLysAenCysLysGlyLysValSerPheThrAaspAenValAlaSerCysGlyGlyG1 402  
 QY 4186 AGTGGTTTAAAGGCATTTGCTTTTCAAGCAATGAAGGAGGATATTTCTTCGAGG 4245

Db	402	yValValTyrLyGlyThrValLeuPheLysAspAsnGluGlyGlyLeuPheArgG1	422
Qy	4246	GAACACAGCATACGATGATTTAAAGGATTCTTGCTGTACTAATAACAGCATCAGAAATACGGA	4305
Db	422	YAsnThrAlaTyrAspAspLeuGlyLeuAlaThrSerArgAspGlnAAsnThrG1	442
Qy	4306	GACAGAGCGGTGAGGAGTATTCTCTCCAGATGATCTGTAAAGTTTGAAGCAA	4365
Db	442	uThrGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAs	462
Qy	4366	TAAAGTTCATTCTTTTGATTACAACTTTGCAAAAGCGAGGCGGAAGCATCTCTAAC	4425
Db	462	nLysGlySerIleValPheAspTyrAsnPheAlaLysGlyArgGlySerIleLeuTh	482
Qy	4426	GAAGAATTC 4435	
Db	482	rLysGluPhe 485	
RESULT 38			
AAU38901			
XX	AAU38901	standard; protein; 1016 AA.	
XX	AAU38901;		
XX	11-SEP-2003	(revised)	
DT	16-JAN-2002	(first entry)	
XX	C. trachomatis CT872 protein.		
XX	Chlamydia; sexually transmitted disease; PID; antibacterial;		
KW	pelvic inflammatory disease; antigen; trachoma; gynecological;		
KW	acute respiratory tract infection; atherosclerosis; male infertility;		
KW	coronary heart disease.		
XX	Chlamydia trachomatis; serovar D.		
OS	W0200181379-A2.		
PN	01-NOV-2001.		
XX	23-APR-2001; 2001WO-US013081.		
XX	21-APR-2000; 2000US-0198853P.		
PR	20-JUL-2000; 2000US-0219752P.		
PR	(CORI-) CORIXA CORP.		
XX	Bhatia A, Probst P, Stromberg EJ;		
XX	WPI; 2001-616771/71.		
DR	N-FSDB; AAS57033.		
DR	New polynucleotide for treating Chlamydia infections encodes a		
PT	polynucleotides containing an immunogenic portion of a Chlamydia antigen.		
PT	Claim 13; Page 171-173; 208pp; English.		
PS	The invention relates to isolated polynucleotide encoding at least a		
XX	partial Chlamydia protein which is an antigenic fragment, or the		
CC	complements, fragments, homologues and variants, and antibodies raised		
CC	against the antigenic proteins (or fragments). The nucleic acids,		
CC	proteins and antibodies are used to diagnose and treat Chlamydia		
CC	infections (e.g. a sexually transmitted disease, pelvic inflammatory		
CC	disease (PID), acute respiratory tract infection, trachoma,		
CC	atherosclerosis and coronary heart disease) in a patient, and in the		
CC	treatment of male infertility. The compounds of the invention are also		
CC	useful for detecting the presence of Chlamydia in a patient, and		
CC	stimulating and/or expanding T cells specific for a Chlamydia protein.		
CC	The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-		
CC	2003 to standardise OS field)		
XX	Sequence 1016 AA;		
SQ			







**THIS PAGE BLANK (USPTO)**



Score: 5131.50 Matches: 983  
Percent Similarity: 98.8% Conservative: 18  
Best Local Similarity: 97.0% Mismatches: 11  
Query Match: 65.1% Indels: 1  
DB: 11 Gaps: 1

US-10-701-844-1 (1-4435) x US-11-103-957-9 (1-1013)

```
Qy 382 ATGCAAGCTCTTCCATAAGTCTCTTCTTCAATGATTTAGCTTATTTCTGCTGCT 441
Db 1 MetGlnThrSerPheHisLeuPheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
Qy 442 TTAATGGGGGGGATATGACGAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
Qy 502 ACGTTAACTGATCATTTCCCTATCTATAGGATCGAGTGGGACTGTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTGAGGAGAGTTAAACATTAAAAATCTTGACAAATCTATTGACAGCTTTGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Qy 622 TGTGTTGGAACTTATTAGGAGTTTACTGTTTGGGAGGAGACACTGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy 742 TTTACTATTGAGGGTTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTACTTCCGTA 801
Db 121 PheThrIleGluGlyPheLeuLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy 802 CTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACAACTCTACACCGTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
Qy 862 AATGGTACTATTATTCTTAAACAGACTTTTGTGTACTCAATATGAGAAGTTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLeuSerLeuThrValGln 200
Qy 982 GGAATTAGCAAGCTTTGCTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy 1042 CAAGTAGTACCAGTTCTCTGCTATGCTTAACGAGGCTCTATTGCTTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Qy 1102 GTTGAGGAGTAAGAGGGGGAGGATTCGTGTTTCAGGATGGGAGGAGGAGTGTCAT 1161
Db 241 ValAlaGlyValArgGlyGlyIleAlaValGlnAspGlyGlnGlnGlyValSer 260
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATATCTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
Qy 1222 GATGGGACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAAAGTTGCTTCTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
Qy 1282 AATATGGAAAACCTTGTCTCAACAAATGTGCTTCTCTGTTTACATTTGCTGCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
Qy 1342 CAACCAACAGTGGACAGGCTTCTAATACGAGTAATTAATACGAGATGGAGGACTATC 1401
```

```
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
Qy 1402 TTCTGTGAAGATGTCGCGCAA--GCAGGATCCATAAATCTCTGATCAGTTCTCTTGTAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
Qy 1459 GGAGAGGAGTAGTTTCTTTTAGTAGCAATGTAGTCTGCTGGGAAAGGGGAGCTATTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380
Qy 1519 GCCAAAAGCTCTCGTGTCTAATCTGTGGCCCTGCTCAATTTTAAAGGAATATCGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
Qy 1579 GATCGTGAGCGATTTATTAGGAGAAATCTCGAGAGCTCAGTTTATCTGCTGATTTATGA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGlyLeuSerGlyGlyLeuSerLeuSerAlaAspTyrGly 420
Qy 1639 GATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAGAATGCTGCGATGTTAAT 1698
Db 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
Qy 1699 GCGTTAACTGTCCTCACAGCCATTTGATGGATCGGGGCGGAGGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
Qy 1759 AGACTTAAAGCAGGGCATCAGATTTCTTTTAATGATCCCATCGAGATGGCAACGAAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
Qy 1819 AACGAGCAGCGAGTCTTCCAACTTCAAAATTTAAACGATGCTGAAGGATACACAGGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
Qy 1879 GATATTGTTTGTCTTAAAGCAGTACTTTGTACCAAAATTTACCAAAATGTTACGATAGCAAGA 1938
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
Qy 1939 AGGATTGTTCTTCGTAAGGCGAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
Qy 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACACACCACCAAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
Qy 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCAATTTGCTCTTCTTCTTTG 2118
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
Qy 2119 TTAGCAACAATGAGTATTACGAATCTCTACCAATCTCCAGCGCAAGATTTCTCATCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
Qy 2179 GCAGTCAATGTGTAGCAACACTGCTGTTCTGTTTACAAATTAGTGGGCTTATCTTTTTCAG 2238
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Qy 2239 GATTGGAATGATACAGCTTATGATAGGTATCATGGCTAGGTTCCTTAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
Qy 2299 GTCTCGAAATACAGTTAGGAGCTAAGCCCCCAGCTAATGCCCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
Qy 2359 GGGAAATGAGATGCTTAAGTATGGCTATCAAGGAGCTGGAGCTTGGCTGGATCTCTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680
Qy 2419 ACAGCAATAATGTCCTTACTCTGAAAGCTACATGCTATAAATCTGGGTATATCTCT 2478
```

```
Db      681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
Qy      2479 GGGCTTAGCGAGTAGCTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538
Db      701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
Qy      2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCCAGGATTA 2598
Db      721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Qy      2599 TGGGTTCTGGAGTTTCGAATTTCTTATCATGACCGGATGCTTTTAGTCAGGATAT 2658
Db      741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy      2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTACTTTGGATCATCGATGTT 2718
Db      761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerMetPhe 780
Qy      2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGTTCCAAAT 2778
Db      781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy      2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTGCATCTAT 2838
Db      801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy      2839 TTGTTCCGAGATCGGTTTATCCGTCTAGCTAGCTAGCGGTTTGGGAATCAGCATATGAAC 2898
Db      821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy      2899 TCATATACATTTGCAGAGAGCGATGTTGTTGGGATAAATCTGCTGGCTGGAGAG 2958
Db      841 SerTyrThrPheAlaGluGluSerAspValArgTrpAsnAsnCysLeuValGlyGlu 860
Qy      2959 ATTGAGCGGGATTACCGATTGTGATTCTTCCATCTAAGCTCTTATTGAATGAGTTGCGT 3018
Db      861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy      3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGACGAT 3078
Db      881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy      3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCAGTCTCTGTGGAGTGAAG 3138
Db      901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy      3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGGCGGCTTATCTGT 3198
Db      921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy      3199 GATGCTTATCGCAACCATCTGTGTACTGAGACAAACGCTCTCTATCCCATCAAGAGATGG 3258
Db      941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy      3259 ACNACAGATGCCCTTTCATTTAGCAAGCATGGAGTTGTGGTTAGAGATCTATGATGCT 3318
Db      961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Qy      3319 TCTCTAACAGTAAATAGAGTATATGGCCATGGAAGATATCAGATTCAGATGCTTCT 3378
Db      981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy      3379 CGAGGCTATGGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db      1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

## RESULT 2

```
US-11-018-868-21
; Sequence 21, Application US/1101868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
```

```
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018.868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-21
```

## Alignment Scores:

Pred. No.:	0	Length:	1013
Score:	5131.50	Matches:	983
Percent Similarity:	99.8%	Conservative:	18
Best Local Similarity:	97.0%	Mismatches:	11
Query Match:	65.1%	Indels:	1
DB:	11	Gaps:	1

US-10-701-844-1 (1-4435) x US-11-018-868-21 (1-1013)

Qy	382	ATGCAAAAGCTCTTTCCATAAGTTCTTTCTTTCAATGATTCTAGCTTATTCTTCTCTCT	441
Db	1	MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer	20
Qy	442	TTAAATGGGGGGATATGCAGCAAAATCATGTTTCTTCAAGGAATTTACGATGGGAG	501
Db	21	LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu	40
Qy	502	ACGTTAACTGTATCTATTTCCCTATCTCTTATAGAGATCCGAGTGGGACTCTGTTT	561
Db	41	ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe	60
Qy	562	TCGCGAGGAGAGTTAAACATTAATAAATCTTGCAATTTCTTATTCAGCTTTGCTTAAAGT	621
Db	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
Qy	622	TGTTTGGAACTTATTAGGAGTTTTACTGTGTTTATAGGAGGAGGACACTCGTTGACTTTC	681
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
Qy	682	GAGAACATGAGACTTCTACAAATGGGCGAGCTCTTAAGTAATAGCGCTGCTGATGACATG	741
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
Qy	742	TTTACTATTGAGGTTTTAAAGAATTTATCTTTTCCAAATTTCAATTCATTACTTGCCGTA	801
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
Qy	802	CTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACAAATCACTACACCGTCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer	160
Qy	862	AATGGTACTATTATTCTAAAAACAGATCTTTGTTACTCAATATCAGAGTTCTCATTC	921
Db	161	AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnGluLysPheSerPhe	180
Qy	922	TATAGTAATTTAGTCTCTCGAGATGGGAGCTATAGATGCTAAGAGCTTTAAACGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln	200
Qy	982	GGAAATTAGCAAGCTTTGTGCTTCTCCAGAAATATCTGCTCAAGCTGATGGGAGCTTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
Qy	1042	CAAGTAGTCACCAAGTTTCTCTGCTATGGCTTAACGAGGCTCTTATTCCTTTGAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240







Query Match:	20.2%		Indels:	15
DB:	11		Gaps:	2

US-10-701-844-1 (1-4435) x US-11-018-868-22 (1-1016)

QY	3451	ATGCCCTTTTCTTGAGACTACATCAATTTGTTTTTAGCTGTGGTTCCTATTCCG	3510
Db	1	MetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCysEryrSer	20
QY	3511	TATGGATTCCGAGCTCTCCTCAAGTGTAACGCCTAATGTAAACACCTCTTTTAAAGGA	3570
Db	21	TyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrProPheLysGly	40
QY	3571	GACGATGTTTACTTGAATGGAGATCGCGCTTTTGTCAAATGTTATGCGAGGAGCTGAAGA	3630
Db	41	AspAspValTyrLeuAlaSerSerProGlnValLeuThrProAsnValTyrAlaGluAsn	60
QY	3631	GGTTCGATTATCTCACTAATGGCGCAATTTAAAGATTACCAGGACAACCATACATTA	3690
Db	61	GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu	80
QY	3691	TCATTACAGATTCTCAAGGCCAGTCTTCTCAAAATTATGCTTCATTTACGACGAGAG	3750
Db	81	SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu	100
QY	3751	ACATTACTCTGAGAGATTTTTCAGTCTGATGTTCTCGAAAAATGTTCTTCGGGAGAA	3810
Db	101	ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu	120
QY	3811	AAGGAATGATCTCCGGAAAACCGTATTTCCGGAGCAGCGAAGTGAATTTCTGG	3870
Db	121	LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyValIlePheTrp	140
QY	3871	GATAACTCCGTGGGTATCTCTCTTATCTACTGTGCCAACCTCATCAACTCGGCT	3930
Db	141	AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro	160
QY	3931	GCT-----CCACAGTTAGTGATCGCG	3954
Db	161	AlaProAlaProAlaProAlaAlaSerSerSerLeuSerProThrValSerAspAlaArg	180
QY	3955	AAAGGCTATTTTCTGTAGAGACTAGTTTTCGAGACTCTCAGCGCTCAAAAAAGGGTC	4014
Db	181	LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLysGlyVal	200
QY	4015	ATGTTCCATATAATATGCGGGAATTTTCGGAACAGTTTTCGAGTAAGAATAATAATAT	4074
Db	201	MetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsn	220
QY	4075	GCTGGTGGAGGACGTGGTTCGGCTACACCATCACTACGACTTTTACAGTTAAATAA	4134
Db	221	AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLysAs	238
QY	4135	CTGTAAAGGGAAGTTCTTTTCAGATAACGTAGCTCTTTCGGAGGCGGAGTGTTTA	4194
Db	238	nCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyValValty	258
QY	4195	TAAAGGCATTTGTCTTTTCAAGACAATGAAGGAGGCATATTTCTTCGGAGGGAACACAG	4254
Db	258	rLysGlyThrValLeuPheLysAspAsnGluGlyIlePhePheArgGlyAsnThrAl	278
QY	4255	ATACATGATTTAGGATTTCTGTCTACTATCACTACAGTACAGTAACGAGACAGGAG	4314
Db	278	aTyrAspAspLeuGlyIleLeuAlaLaThrSerArgAspGlnAsnThrGluThrGlyGl	298
QY	4315	CGGTGGAGGATTTTGTCTCCAGATGATTTCTGTAAAGTTTGAAGGCAATAAAGGTTTC	4374
Db	298	yGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySe	318
QY	4375	TATGTTTTTGAATACATTTTCAAAAGGACAGGCGGAGGATTCCTAACGAAGATTT	4434
Db	318	IleValPheAspTyrAsnPheAlaLysGlyArgGlySerIleLeuThrLysGluPh	338

QY 982 GGAATTAGCAAGCTTTGTCTCTTCAAGAAATACTGCTCAAGCTGATGGGGAGCTGT 1041  
Db 159 HisAsnTyrLeuPheThrAlaPheGluGluAsnSerSerLysGlyAsnGlyAlaIle 178  
QY 1042 CAAGTAGTCACCAAGTTCTCTCTATGCTTAACAGAGCTCTTATTCCTTTAGCGAAT 1101  
Db 179 GlnAlaGlnThr--PheSerLeuSerArgAsnValSerProIleSerPheAlaArgAsn 197  
QY 1102 GTTGAGGAGTAAGAGGGAGGATGCTGCTGTTTCAGGATGGCGAGGAGTGCA 1161  
Db 198 ArgAlaAspLeuAsnGlyAlaIleCysCys-- 208  
QY 1162 TCATCTACTTCAACAGAGTCCAGTAGTAAGTTTTCAGAAATACTCGGCTAGAGTTT 1221  
Db 208 ----- 208  
QY 1222 GATGGAACTAGCCGAGTAGGAGGAGTTTACTCTACGGGAACGTTGCTTTCCTG 1281  
Db 209 -----SerAsnLeuIleCysSerGlyAsnVal----- 217  
QY 1282 AATAATGGAAAAACCTGTTTCTCAACATGTTGCTTCTCTGTTTACATTCGCTAG 1341  
Db 218 -----AsnProLeuPhe----- 222  
QY 1342 CAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATTACGAGATGGAGGACTATC 1401  
Db 223 -----ThrGlyAsnSerAlaThrAsnGlyAlaIle 233  
QY 1402 TTCTGTGAAGATGTCGCAAGCAGGATCCAACTCTGGATCAGTTTCTTTGATGGA 1461  
Db 234 CysCysIle-----SerAspLeuAsnThrSerGluLysGlySerLeuAlaCys 251  
QY 1462 GAGGGAGTAGTTTCTTTAGTAGCAATGATGCTGCGGAGGGAGCTATTTATGCC 1521  
Db 252 AsnGlnGluThrLeuPheAlaSerAsnSerAlaLysGluLysGlyAlaIleTyrAla 271  
QY 1522 AAAAGCTCTCGTTGCTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1581  
Db 272 LysHisMetValLeuArgTyrAsnGlyProValSerPheIleAsnAsnSerAlaLysIle 291  
QY 1582 GGTGAGGCAATTTATTAGGAGATCTGAGAGCTCAGTTTATCTGCTGATTTAGGAGAT 1641  
Db 292 GlyGlyAlaIleAlaIleGlnSerGlyGlySerLeuSerIleLeuAlaGlyGluGlySer 311  
QY 1642 ATTATTTTCAGTGGGAATCTTAAAGAACAGCAAGAGAGATCTCCGATGTTATGCG 1701  
Db 312 ValLeuPheGlnAsnAsnSerGlnArgThrSerAspGlnGlyLeuValArg---AsnAla 330  
QY 1702 GTAACGTGTCCTCACAGCCATTTCGATGGGATCGGAGGAAATAACGACATTAAAGA 1761  
Db 331 IleTyrLeuGluLysAspAlaIle-----LeuSerSerLeuGlu 343  
QY 1762 GCTAAAGCAGGCGATCAGATTTCTTTATGATGCCATC---GAGATGGCAACAGGAAAT 1818  
Db 344 AlaArgAsnGly---AspIleLeuPheAspProIleValGlnGluSerSerLys 362  
QY 1819 AACCAAGCAGGCGAGTCTTCCAACTCTTAAATAATCAAGATGGTGAAGATCACAGGG 1878  
Db 363 GluSerProLeuProSerSerLeuGlnAlaSerValThrSerProThrProAlaThrAla 382  
QY 1879 -----GATATTGTTTGTCT----- 1893  
Db 383 SerProLeuValIleGlnThrSerAlaAsnArgSerValIlePheSerSerGluArgLeu 402  
QY 1894 -----AATGGAGCAGTACTTTGTACCAAAAATGTTTACGATA 1929  
Db 403 SerGluGluLysThrProAspAsnLeuThrSerGlnLeuGlnProIleGluLeu 422  
QY 1930 GAGCAAGGAGGATTTGTTCTTCGTGAAAGAGCAAAATATTCAGTGAATTTCTTAAGTCAG 1989  
Db 423 LysSerGlyArgLeuValLeuLysAspArgAlaValLeuSerAlaProSerLeuSerGln 442  
QY 1990 ACAGGTGGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAA 2046

Db 443 AspProGlnAlaLeuLeuIleMetGluAlaGlyThrSer----- 455  
QY 2047 CCACCAACAGCCTCTCCGCGCTAATCAGTTGATCAGCTTCCCAATCTGATTTGCT 2106  
Db 456 -----LeuLysThrSerSerAspLeuLysLeuAla 465  
QY 2107 CTTTCTTCTTTGTAGCAAAACAATGCAGTTACGAATCTCTTACCAATCTCTCCAGCGCAA 2166  
Db 466 ThrLeuSerIleProLeuHisSerLeu----- 474  
QY 2167 GATTCTCATCTCGCAGTCATTCGTAGCAACAACGCTGTTCTGTTTACAAATTAGTGGCCT 2226  
Db 475 -----AspThrGluLysSerValThrIleHisAlaPro 485  
QY 2227 -----ATCTTTTGTAGGATTTG---GATGATACAGCTTATGATGG 2265  
Db 486 AsnLeuSerIleGlnLysIlePheLeuSerAsnSerGlyAspIleAsnPheTyrGluAsn 505  
QY 2266 TATGATTGGCTAGGTTCTAATCAAAAAATCAATGCTCCTGAAATTTACAGTTAGGAGCTAAG 2325  
Db 506 ValGluLeuLeuSerLysGluGlnAsn---AsnIleProLeuLeuThrLeu----- 521  
QY 2326 CCCCAGCTAATGCCCATCAGATTGACTCTA-----GGGAATCAGATGCTAAGTAT 2379  
Db 522 -----SerLysGluGlnSerHisLeuLeuProAspGlyAsnLeuSerSerHisPhe 539  
QY 2380 GGTATCAAGAGCTGGAAGCTTCGCTGGGATCTTAATACAGCAATAATGCTCTTAT 2439  
Db 540 GlyTyrGlnGlyAspTrpThrPheSerTrp-----LysAspSerAspGluGly---His 556  
QY 2440 ACTCTGAAAGCTACATGGAATAAACTGGGTATAATCTCTGGCTCAGCAGTAGCTTCT 2499  
Db 557 SerLeuIleAlaAsnTrpThrProLysAsnTyrValProHisProGluArgGlnSerThr 576  
QY 2500 TTGGTTTCCAAATAGTTTATGGGATCCATTTTAGATATACGATCTCGCATTCAGCAATT 2559  
Db 577 LeuValAlaAsnThrLeuTrpAsnThrTyrSerAspMetGlnAlaValGlnSerMetIle 596  
QY 2560 CAAGCAGTGTGATGGGCGCTTATTCGAGGATTTATGGTTCGAGGATTCGATTCGAAT 2619  
Db 597 AsnThrIleAlaHisGlyGlyAlaTyrLeuPheGlyThrTrpGlySerAlaValSerAsn 616  
QY 2620 TTCTCTCAT---CATCACCGCATGCTTTAGGT---CAGGATATCGGTATATAGTGG 2673  
Db 617 LeuPheTyrAlaHisAspSerSerGlyLysProIleAsnTrpHisHisArgSerLeu 636  
QY 2674 GGTATTCCTTAGGACCAACTCTTAC---TTTGGATCATCGATGTTGCTTAGCAATT 2730  
Db 637 GlyTyrLeuPheGlyIleSerThrHisSerLeuAspAspHisSerPheCysLeuAlaAla 656  
QY 2731 ACCGAAGTATTTGCTAGATCTAAAGATTATGATGTCGTTCGTTCCCAATCATCATGCTGCG 2790  
Db 657 GlyGlnLeuGlyLysSerAspSerPheIleThrSerThrGlu----- 672  
QY 2791 ATAGGATCGTTTATCTATCTACCAACAGCTTTTATGTGATCCTTATTTGTTCCGAGAT 2850  
Db 673 ---ThrThrSerTyrIleAlaThrValGlnAlaGlnLeuAlaThrProLeuMet---Lys 690  
QY 2851 GCGTTTATCGTCTAGCTACCGGTTTGGGAATFCAGCATATGAAACCTCATAT---ACA 2907  
Db 691 IleSerAlaGlnAlaCysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSer 710  
QY 2908 TTTGCAGAGGAGGAGCGATGTTTCGTTGGGATAATACTGCTGCTGCGAGAGATTGGAGCG 2967  
Db 711 PheSerLysGluGlyPheGlySerTrpHisSerValAlaValSerGlyGluValCysAla 730  
QY 2968 GGATTACCGATCTGATTACTCCATCTAAGCTCTTATTTGAATGAGTTGCTGCTTCTGCTG 3027  
Db 731 SerIleProIleValSerAsnGlySerGlyLeuPhe---SerSerPheSerIlePheSer 749  
QY 3028 CAAGCTGAGTTTCTTATGCTCGCATCATGATCTTTTACAGAGGAAGCGCATCAAGCTCGG 3087  
Db 3087



1819 ACCAGCAGCGCAGCTCTTCCAACTTCTTAAATAATACAGTGTGAGGATACACAGG 1878  
363 GluSerProLeuProSerSerLeuGlnAlaSerValThrSerProThrProAlaThrAla 382  
1879 -----GATATTGTTTTCCT----- 1893  
383 SerProLeuValIleGlnThrSerAlaAsnArgSerValIlePheSerSerGluArgLeu 402  
1894 -----AATGAACAGTACTTTGTACCAAAATGTTAGGATA 1929  
403 SerGluGluGluThrProAspAsnLeuThrSerGlnLeuGlnProIleGluLeu 422  
1930 GAGCAAGGAGGATGTTCTCGTGAAGGAGGAAATATCAGTGAATCTCTAAGTCAG 1989  
423 LysSerGlyArgLeuValLeuLysAspArgAlaValLeuSerAlaProSerLeuSerGln 442  
1990 ACAGGTGGGAGT---CTGTATATGAGGCTGGGAGTACATGGGATTTTGTAACTCCCAA 2046  
443 AspProGlnAlaLeuLeuMetGluAlaGlyThrSer----- 455  
2047 CCACCAACAGCCTCTCGCGCTAATCAGTTGATCAGCGTTTCCAAATCTGCAATTTGTCT 2106  
456 -----LeuLysThrSerSerAspLeuLysLeuAla 465  
2107 CTTTCTTCTTTGTTAGCAACAAATGCGATTACGAATCCTCTACCAATCCTCCAGCGCAA 2166  
466 ThrLeuSerIleProLeuHisSerLeu----- 474  
2167 GATTCTCATCTCGAGTCATGTTAGCACAACCTGCTGTTCTGTTACAAATTTAGTGGCCT 2226  
475 -----AspThrGluLysSerValThrIleHisAlaPro 485  
2227 -----ATCTTTTGTAGGATTG---GATGATCAGCTTATGATAGG 2265  
486 AsnLeuSerIleGlnLysIlePheLeuSerAsnSerGlyAspGluAsnPheThrGluAsn 505  
2266 TATGATTGCTAGGTTCTAATCAAAAATCAATGCTCTGAAATTTACAGTTAGGCACTAAG 2325  
506 ValGluLeuLeuSerLysGluGlnAsn---AsnIleProLeuLeuThrLeu----- 521  
2326 CCCCAGCTAATGCCCATCAGATTGACTCTA-----GGGAATGAGATGCCCTAAGTAT 2379  
522 -----SerLysGluGlnSerHisLeuHisLeuProAspGlyAsnLeuSerHisPhe 539  
2380 GGCTATCAAGAACTGGAAGCTTGGTGGGATCCTTAATACAGAAATAATGTCCTTAT 2439  
540 GlyTyrGlnGlyAspThrPheSerTrp-----LysAspSerAspGluGly---His 556  
2440 ACTCTGAAGCTACATGAGCTAAACCTGGGTATATCTGGGCTGAGCGAGTAGCTTCT 2499  
557 SerLeuIleAlaAsnTrpThrProLysAsnTrpValProHisProGluArgGlnSerThr 576  
2500 TTGGTTCCAAATAGTTTATGGGATCCATTTTATAGATATAGCTGGCATTCCAGCAAT 2559  
577 LeuValAlaAsnTrpLeuTrpAsnThrTrpSerAspMetGlnAlaValGlnSerMetIle 596  
2560 CAAGCAAGTGTGGATGGCGCTCTTATTTGCGAGGATATGGGTTTCTGGAGTTTCAAT 2619  
597 AsnThrIleAlaHisGlyGlyAlaTyrLeuPheGlyThrTrpGlySerAlaValSerAsn 616  
2620 TTCTTCTAT---CATGACCGGATGCTTTAGT---CAGGATATCGGTATATTTAGTGG 2673  
617 LeuPheTyrAlaHisAspSerSerGlyLysPheProIleAspAsnTrpHisHisArgSerLeu 636  
2674 GGTATTCTCTTAGGAGCAAACTCTAC---TTTGGATCATCGATGTTTGGCTAGCAATTT 2730  
637 GlyTyrLeuPheGlyIleSerThrHisSerLeuAspAspHisSerPheCysLeuAlaAla 656  
2731 ACCGAAGTATTTGGTAGATCTAAAGATTTATGTAGTGTGCTGCTCCAAATCATGCTTGC 2790  
657 GlyGlnLeuLeuGlyLysSerSerAspPheIleThrSerThrGlu----- 672  
2791 ATAGATCCGTTTATCTATCTACCAACAAGCTTTATGTGGATCCTATTGTTCCGGAGAT 2850

673 ---ThrThrSerTyrIleAlaThrValGlnAlaGlnLeuAlaThrProLeuMet---Lys 690  
2851 GCGTTTATCCGTCGTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATAT---ACA 2907  
691 IleSerAlaGlnAlaCysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSer 710  
2908 TTTGACAGAGGAGCGGATGTTCTGGTGGATATAACTGTCTGGCTGGAGAGATTGGAGCG 2967  
711 PheSerLysGluGlyPheGlySerTrpHisSerValAlaValSerGlyGluValCysAla 730  
2968 GGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGATGATGAGTTGGCTCTTTCGCG 3027  
731 SerIleProIleValSerAsnGlySerGlyLeuPhe---SerSerPheSerIlePheSer 749  
3028 CAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAGCGGATCAAGCTCG 3087  
750 LysLeuGlnGlyPheSerGlyThrGlnAspGlyPheGluGluSerSerGlyGluLeuArg 769  
3088 GCATTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTTGATCGA 3147  
770 SerPheSerAlaSerSerPheArgAsnIleSerLeuProMetGlyIleThrPheGluLys 789  
3148 TGTCTAGTACATCTCTAATAATATAGCTTTTATGGCGCTTATATCTGTGATGCTTAT 3207  
790 LysSerGlnLysThrArgAsnTyrTyrPheLeuGlyAlaTyrIleGlnAspLeuLys 809  
3208 CCGACATCTCTGGTACTGAGACAAGCTCTCTATCCCATCAAGAGACATGGACACAGAT 3267  
810 ArgAspValGluSerGlyProValValLeuLeuLysAsnAlaValSerTrpAspAlaPro 829  
3268 GCCTTTTCTATTAGCAAGACATCGAGTTGTGGTGTAGAGGATCTATGTATGCTTCTTAACA 3327  
830 MetAlaAsnLeuAspSerArgAlaTyrMetPheArgLeuThrAsnGlnArgAlaLeu--- 848  
3328 AGTAATATAGATATATGCGCATGAGCAAGATGAGATGAGATGAGATGCTTCTCGAGGCTAT 3387  
849 HisArgLeuGlnThrLeuLeuAsnValSerTyrValLeuArgGlyGlnSerHisSerTyr 868  
3388 GGTTCAGTGCAGGAGTAGTCCGGTTC 3417  
869 SerLeuAspLeuGlyThrThrTyrArgPhe 878

RESULT 7  
US-11-103-957-13  
; Sequence 13, Application US/11103957  
; Publication No. US20050281847A1  
; GENERAL INFORMATION:  
; APPLICANT: Berthet, Francois-Xavier Jacques  
; APPLICANT: Lobet, Yves  
; APPLICANT: Poolman, Jan  
; APPLICANT: Verlant, Vincent Georges Christian Louis  
; TITLE OF INVENTION: Vaccine Composition  
; FILE REFERENCE: B45261  
; CURRENT APPLICATION NUMBER: US/11/103,957  
; PRIOR FILING DATE: 2005-04-12  
; PRIOR APPLICATION NUMBER: US/10/467,534  
; PRIOR FILING DATE: 2004-02-03  
; PRIOR APPLICATION NUMBER: PCT/EP02/01356  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: GB 0103169.9  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 964  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-11-103-957-13  
Alignment Scores: 5.88e-48 Length: 964  
Pred. No.: 649.50 Matches: 260  
Score:





Db 255 AenArgGlyAenIleValPheTyrAenAen-----ArgCysPheLysAen----- 269  
QY 1318 TCTCCTGTTTACATTGCTGCTAAAGCAACCAAGTGGACAGGCTCTTAATACGAGTAAT 1377  
Db 270 -----ValGluThrAlaSerSer 275  
QY 1378 AATTACGAGATGGAGGAGCTATCTTCTGTAAGAAATGCTGGCAAGCAGGATCCATTAAC 1437  
Db 276 GluAlaSerAspGlyAlaIleTyrValThrArgLeuAaspValThrGlyAen--- 294  
QY 1438 TCTGGATCAGTTCTCTTGGATGGAGGAGTAGTTTCTTCTAGTAGCAATGCTAGCTGCT 1497  
Db 295 -----ArgGlyArgIlePhePheSerAspAenIleThrLys 306  
QY 1498 GGGAAAGGGGAGCTATTATGCCAAAAGCTCTCGTTGCTAACTGTGGCCCTGTACAA 1557  
Db 307 AenTyrGlyGlyAlaIleTyrAlaProValValThrLeuValAaspAenGlyProThrTyr 326  
QY 1558 TTTTAAAGAAATATCGCTAATGAT---GGTGGAGGATTTATTTAGGAGAAATCTGGAGAG 1614  
Db 327 PheIleAenAenIleAlaAenAenLysGlyAlaIleTyrIleAaspGlyThrSerAen 346  
QY 1615 CTCAGTTTATCTGCTGATATGGAGATATTATTTTCGATGGGAATCTTAAAGAACAGCC 1674  
Db 347 SerLysIleSerAlaAaspArgHisAlaIleIlePheAenGluAenIle----- 362  
QY 1675 AAAGAGAAATGCTCGCGATGTTAATGCGTAACGTGCTCTCA-----CAA 1719  
Db 363 ValThrAenValThrAenAlaAenGlyThrSerThrSerAlaAenProProArgAen 382  
QY 1720 GCCATTCGATGGGATCGGAGGGAATAACGACATTAAGAGCTAAAGCAGGCGATCAG 1779  
Db 383 AlaIleThrValAlaSerSerSerGlyGluIleLeuLeuGlyAlaGlySerSerGlnAen 402  
QY 1780 ATTCTCTTTAATGATCCCATCGAGATGGCAACGGAATAACACCGCAGCGCTCTCC 1839  
Db 403 LeuIlePheTyrAaspProIleGluValSerAen-----AlaGly 415  
QY 1840 AAATCTCTAAAATTAACGATGGTGAAGGATACACAGGGATATTGTTTTCCTAAT--- 1896  
Db 416 ValSerValSerPheAenLysGluAlaAaspGlnThrGlySerValValPheSerGlyAla 435  
QY 1897 -----GGAGCAGTACTTTGTACCAAAAT-----GTT 1923  
Db 436 ThrValAenSerAlaAaspPheHisGlnArgAenLeuGlnThrLysThrProAlaProLeu 455  
QY 1924 ACGATAGACAGGAGGATGCTCTCTGTAAGGCAAAATTAATCAGTGAATCTCTA 1983  
Db 456 ThrLeuSerAenGlyPheLeuCysIleGluAaspHisAlaGlnLeuThrValAenArgPhe 475  
QY 1984 AGTCAGACAGGTTGGAGTCTGTATATGGAAGCTGGAGT----- 2022  
Db 476 ThrGlnThrGlyValValSerLeuGlyAenGlyAlaValLeuSerCysTyrLysAen 495  
QY 2023 ---ACATGGGATTTGTAATCTCCACACACCACAGCCCTCTGCGCTAATCAGTTG 2079  
Db 496 GlyThrGlyAasp-----SerAlaSerAenAlaSer 505  
QY 2080 ATCAGCTTTCCAAATCGATTGCTCTCTTCTTTGTTAGCAACAATCA----- 2133  
Db 506 IleThrLeuLysHisIleGlyLeuAenLeuSerSerIleLeuLysSerGlyAlaGluIle 525  
QY 2134 -----GTTACGAATCTCTACCAATCTCCAGCGCAAGATTCATCCTCGAGTCATT 2187  
Db 526 ProLeuLeuTipValGluProThrAen-----AenSerAenAenTyrThrAlaAaspThr 543  
QY 2188 GTAGCACA---ACTGCTGGTCTGTACAAATAGTGGCCCTATCTTTTTCAGGATTTG 2244  
Db 544 AlaAlaThrPheSerLeuSerAaspValLysLeuSer-----LeuIleAaspTyr 560  
QY 2245 GATGATACAGCTTATGATAGGTATGATTGGCTAGGTTCTTAATCAAAAATCAATGCTCTG 2304  
Db 2245 ----- 2304

Db 561 GlyAenSerProTyrGluSerThrAasp-----Leu 570  
QY 2305 AANTTACAGTTAGGAGCTAAGCCC-----CCAGCTAATGCCCATCAGATTTGACTCTA 2358  
Db 571 ThrHisAlaLeuSerSerGlnProMetLeuSerIleSerGluAlaSerAaspAenGlnLeu 590  
QY 2359 GGGAAATGAG-----ATGCTTAAGTATGCTGCTATCAAGGAAGC 2394  
Db 591 GlnSerGluAenIleAaspPheSerGlyLeuAenValProHisTyrGlyTyrGlnGlyLeu 610  
QY 2395 TGGAAAGCTTGGCTGG----- 2409  
Db 611 TrpThrTrpGlyTyrAlaLysThrGlnAaspProGluProAlaSerSerAlaThrIleThr 630  
QY 2410 GATCTTAATACAGCAATAATAGTCTTATCTCTGAAAGCTACATGCACTAATAAAGCTGG 2469  
Db 631 AaspProGlnLysAlaAenArgPheHisArgThrLeuLeuLeuThrTrpLeuProAlaGly 650  
QY 2470 TATAATCTGGGCTGAGCGAGTAGTCTTCTTCCAAATAGTTTATGGGATCC--- 2526  
Db 651 TyrValProSerProLysHisArgSerProLeuIleAlaAenThrLeuTrpLysAenMet 670  
QY 2527 ATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTAT 2586  
Db 671 LeuLeuAlaThrGluSerLeuLysAenSerAlaGluLeuThrProSerGlyHisProPhe 690  
QY 2587 TGTGAGGATTTAGGTTTCTGGAGTTTCGAATTTCTTCTATCATGACCGCGATCTTTA 2646  
Db 691 Trp---GlyIleThrGlyGlyLeuGlyMetMetValTyrGlnAaspProArgGluAen 709  
QY 2647 GGTGAGGATATCGTATATAGTGGGTTTATCTCTTAGGA-----GCAAACTCC 2697  
Db 710 HisProGlyPheHisMetArgSerSerGlyTyrSerAlaGlyMetIleAlaGlyGlnThr 729  
QY 2698 TACTTTGATCATCGATGTTTGGTCTAGCATTTACGAAGTATTTGGTAGA---TCTAAA 2754  
Db 730 HisThrPheSerLeuLysPheSerGlnThrTyrLysLeuAenGluArgTyrAlaLys 749  
QY 2755 GATTATGATGTTGCTGTTCCATCATCATGCTGTCATGATGATCGGTTTATCTATCTACC 2814  
Db 750 AenAenVal-----SerSerLysAenTyrSerCysGlnGlyGluMetLeuPheSerLeu 767  
QY 2815 CAACAAGCT-----TTATGTCGATCTCTATTTGTCGGAGATGCGTTTATC 2859  
Db 768 GlnGluGlyPheLeuLeuThrLysLeuValGlyLeuTyrSerTyrGlyAaspHisAenCys 787  
QY 2860 CGTGTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGACAGAGGAG 2919  
Db 788 HisHisPheTyrThrGlnGly---GluAenLeuThrSerGlnGlyThrPheArgSerGln 806  
QY 2920 AGCGATGTTGTTGGGATTAATACTGCTGCTGGAGAGATTTGGAGCGGATTTACGATT 2979  
Db 807 Thr-----MetGlyGlyAlaValPhePheAaspLeuProMet 818  
QY 2980 GTGATTACTCCATCAAGCTCTATTTGAATGATTTGCTTCTCGTCAAGCTCAGTTT 3039  
Db 819 LysProPheGlySerThrHisIleLeuThr-----AlaProPheLeuGlyAlaLeuGly 836  
QY 3040 TCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGCGATTCAGAGC 3099  
Db 837 IleTyrSerSerLeuSerHisPheThrGluValGlyAlaTyrProArgSerPheSerThr 856  
QY 3100 ---GGACATCTCTAAATCTATCAGTTCTGTTGGAGTGAAGTTTGTATGATGCTTCTAGT 3156  
Db 857 LysThrProLeuIleAenValProIleGlyValLysGlySerPheMetAenAla 876  
QY 3157 ACACAT---CCTAATAAATATAGCTTTATGGCGCTTATATCTGTGATCTTATCGACC 3213  
Db 877 ThrHisArgProGlnAlaIleThrValGluLeuAlaTyrGlnProValLeuTyrArgGln 896  
QY 3214 ATCTGTTGTTGACAGCAACGCTCTCTATCCCATCAAGAGACATGGAACACAGATGCTTT 3273  
Db 897 GluProGlyIleAlaAlaGlnLeuLeuAlaSerLysGlyIleTyrPheGlySerGlySer 916



```
QY 3274 CATTAGCAAGACATGGAGTCTGCTAGGAGATCTATGATGCTTCT----- 3321
Db 917 ProSerSerArgHisAlaMetSerTyrIysIleSerGlnGlnThrGlnProLeuSerTrp 936

QY 3322 CTAACAAGTAATATAGAGTATATAGC 3348
Db 937 LeuThrLeuHisPheGlnTyrHisGly 945

RESULT 9
US-11-103-957-15
; Sequence 15, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verliant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-15

Alignment Scores:
Pred. No.: 2,52e-43 Length: 1531
Score: 597.50 Matches: 293
Percent Similarity: 36.7% Conservative: 153
Best Local Similarity: 24.1% Mismatches: 474
Query Match: 7.6% Indels: 295
DB: 11 Gaps: 54

US-10-701-844-1 (1-4435) x US-11-103-957-15 (1-1531)
QY 430 TCTTGCTGCTCTTTAAATGGGGGGATATGCAGCAAAATCATGTT----- 477
Db 393 SerThrAlaCysLeuGlyGlyAlaIleAlaAlaGlnGluIleValSerIleGlnAsn 412
QY 478 ---CCTCAAGGAATT---TACGATGGGAGACGTTTAACTGCTATCATTTCCCTATAGTGT 531
Db 413 AsnGlnAlaGlyIleSerPheGluGlyGlyLysAlaSerPheGlyGly----- 428
QY 532 ATAGGAGATCCGAGTGGACTACTGTTTTTCTGCGAGGAGTTAACTTAAATAATCTT 591
Db 429 ---GlyIleAlaCysGlySerPheSerAlaGlyIleAlaSerValLeuGlyThrIle 447
QY 592 GAC-----AATTCATTGACGCTTTCCTTTTAAGT-----TGTTTTGGGAACCTTA 636
Db 448 AspIleSerLysAsnLeuGlyAlaIleSerPheSerArgThrLeuCysThrThrSerAsp 467
QY 637 TTAGGGAGTTTACTGTTTTAGGAGAGGACACTCGTTGACTTCGAGACATACGACT 696
Db 468 LeuGlyGlnMetGluTyrGlnGly---GlyGlyAlaLeuPheGlyGluAsnIleSerLeu 486
QY 697 TCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTAGGGT 756
Db 487 SerGluAsnAlaGlyValLeuThr----- 494
QY 757 TTTAAAGAA-----TTATCTCTTTTCCAATTCGAATTCATTACTTTCGCTACTGCCT 807
Db 495 PheLysAsnIleValLysThrPheAlaSerAsnGlyIleLeuGlyGlyGlyAla 514
```

```
QY 808 GCTGCAACGACACTAATAAGGGTAGCCAGATCCGACGACAAACATCTACACCGTCTAATGGT 867
Db 515 IleLeuAlaThrGlyLysValGluIleThrAsnAsnSerGluGlyIleSerPheThrGly 534
QY 868 ACTATTTATCTAAACACAGACTTTTGTACTCAATAATAGAACTCTCATCTCTATAGT 927
Db 535 Asn-----AlaArgAlaProGlnAlaLeuProThrGlnGluGluPheProLeuPheSer 552
QY 928 AAT-----TTAGTCTCTCGA-----GATGGGGAGCTATAGATGCTAAG 966
Db 553 LysLysGluGlyArgProLeuSerSerGlyTyrSerGlyGlyGlyAlaIleLeuGlyArg 572
QY 967 AGCTTAAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCT 1026
Db 573 GluValAlaIleLeuHis---AsnAlaAlaValPheGluGlnAsnArgLeuGlnCys 591
QY 1027 GAT-----GGGGGAGCTTGTCAAGTAGTAGTCAAC 1053
Db 592 SerGluGluAlaThrLeuLeuGlyCysCysGlyGlyGlyAlaValHisGlyMetAsp 611
QY 1054 AGTTTCTCTGCTATGCTAAACGAGCTCTCTATTGCTTTGTAGCGAATCTTGCA----- 1107
Db 612 SerThrSerIleValcylAsnSerSer---ValArgPheGlyAsnAsnTyrAlaMetGly 630
QY 1108 ---GGAGTAAGAGGGGGGAGTTCGCTGCTGCTTCAGGATGGCGAGCAGGAGTGTCTCA 1164
Db 631 GlnGlyValSerGlyGlyAlaLeuLeuSer----- 640
QY 1165 TCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACTCGCGTAGAGTTTGAT 1224
Db 641 -----LysThrValGlnLeuAlaGlyAsnGlySerValAspPheSer 654
QY 1225 GGGAACTAGCCGAGTAGGAGGAGGATTAC-----TCCTACGGGAGCTGTGCTTTC 1278
Db 655 ArgAsnIleAlaSerLeuGlyGlyAlaLeuGlnAlaSerGluGlyAsnCysGluLeu 674
QY 1279 CTGAATAATAGAAAACCTTG----- 1299
Db 675 ValAspAsnGlyTyrValLeuPheArgAspAsnArgGlyArgValTyrGlyAlaIle 694
QY 1300 -----TTTCTCAAC 1308
Db 695 SerCysLeuArgGlyAspValValIleSerGlyAsnLysGlyArgValGluPheLysAsp 714
QY 1309 AATGTTGCTTCTCTCTTTACATTTGCTGCTAAG-----CAACCA 1347
Db 715 AsnIleAlaThrArgLeuTyrValGluGluThrValGluLysValGluValGluPro 734
QY 1348 ACAAGTGGACAGGCTTCTAATACGAGTAATAATTAC----- 1383
Db 735 AlaProGluGlnLysAspAsnAsnGluLeuSerPheLeuGlyArgAlaGluGlnSerPhe 754
QY 1384 -----GGAGAT----- 1389
Db 755 IleThrAlaAlaAsnGlnAlaLeuPheAlaSerGluAspGlyAspLeuSerProGluSer 774
QY 1390 -----GGAGGAGCTATCTTCTGT 1407
Db 775 SerIleSerSerGluLeuAlaLysArgGluCysAlaGlyGlyAlaIlePheAla 794
QY 1408 AAG-----AATGGTGGCCCAAGCAGGATCCCAATAAC----- 1437
Db 795 LysArgValArgIleValAspAsnGlnGluAlaValPheSerAsnAsnPheSerAsp 814
QY 1438 -----TCTGATCAGTT-----TCCTTTGATGGAGAC 1464
Db 815 IleTyrGlyGlyAlaIlePheThrGlySerLeuArgGluGluAspLysLeuAspGlyGln 834
QY 1465 -----GGAGTAGTTTCTTTAGTCAATGTAGCT 1494
Db 835 IleProGluValLeuIleSerGlyAsnAlaGlyAspValValPheSerGlyAsnSerSer 854
```

QY 1495 GCTGGGAAA-----GGGGAGCTATTATTATGCGCAAAAAGCTCTCG 1533  
Db 855 LysArgAspGluHisLeuProHisThrGlyGlyAlaIleCysThrGlnAsnLeuThr 874  
QY 1534 GTTGCT---AACTGTGGCCCTGACAAATTTTAAAGGAATATCGCTAAATGATGGTGGAGCG 1590  
Db 875 IleSerGlnAsnThrGlyAsnValLeuPheTyrAsnAsnValAlaCysSerGlyGlyAla 894  
QY 1591 ATTTATTATGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATATGAGATATTATTTTC 1650  
Db 895 ValArgIleGluAspHisGlyAsnValLeuGluAlaPheGlyGlyAspIleValPhe 914  
QY 1651 GATGGAACTTAAAGAACGACGCAAGAAAGAAATGCTGCCGATGTTAATGGCGTAACCTGTG 1710  
Db 915 LysGlyAsnSerSerPheArgAlaGln----- 923  
QY 1711 TCCTCACAAAGCATTTCGATG---GGATCGGGAGGGAATAAAGACATTAAAGAGCTAAA 1767  
Db 924 GlySerAspAlaIleTyrPheAlaGlyLysGluSerHisIleThrAlaLeuAsnAlaThr 943  
QY 1768 GCAGGGCATCAGATTCTCTTAATATGATCCCATCGAGATGGCAAAACGGAATAAACACGACA 1827  
Db 944 GluGlyHisAlaIleValPheHisAspAlaLeuValPheGlu-----AsnLeuGluGlu 961  
QY 1828 CGCAGTCTTCCAAACTCTTAAATTAACGATGCGTAA-----GGATACACAGGGGAT 1881  
Db 962 ArgLysSerAlaGluValLeuLeuIleAsnSerArgGluAsnProGlyTyrThrGlySer 981  
QY 1882 ATGTGTTTGTCTAAAGGAGCAGTACTTGTACCAAAATGTTACGATAGACAGAGG 1941  
Db 982 IleArgPheLeuGluAlaGluSerLysValProGlnCysIleHisValGlnGlnGlySer 1001  
QY 1942 ATGTGTTCTTCGAAAGCGCAAAATATCAGTGAATCTCTAAAGTCAG---ACAGGTGGG 1998  
Db 1002 LeuGluLeuLeuAsnGlyAlaThrLeuCysSerTyrGlyPheLysGlnAspAlaGlyAla 1021  
QY 1999 AGTCTGTATATGAAGCTGGAGTACATGGGATTTTGTGA---ACTCCCAACACCAACAA 2055  
Db 1022 LysLeuValLeuAlaGlyAlaLysLysIleLeuAspSerGlyThrProValGln 1041  
QY 2056 CAGCTCTCCGCTCAAT-----CAGTTGATCAGCTTCCCAATCTGCATTG 2103  
Db 1042 GlnGlyHisAlaIleSerLysProGluAlaGluIleGluSerSerSerGluProGluGly 1061  
QY 2104 TCTCTTCTCTTCTGTAGCAACCAATGAGTACGATTCCTCCACCAATCCTCCAGCG 2163  
Db 1062 AlaHisSerLeuTrpIleAlaLysAla-----GlnThrValProMet 1077  
QY 2164 CAAGATTCTCAT----- 2175  
Db 1078 ValAspIleHisThrIleSerValAspLeuAlaSerPheSerSerSerGlnGlnGluGly 1097  
QY 2176 -----CCTGCAGTCAATGGTAGCACACTGCTGGTCTGTTACAAATGATGGG 2223  
Db 1098 ThrValGluAlaProGlnValIle-----ValProGlyGlySerTyrValArgSerGly 1115  
QY 2224 CCACTCTTTTGTAG-----GATTTGGATGATACAGCTTATGATAGGTATGATGGCTA 2277  
Db 1116 GluLeuAsnLeuGluLeuValAsnThrThrGlyThrGlyTyrGluAsnHisAlaLeuLeu 1135  
QY 2278 GGTTCCTAAATAAATCAATCTCTGAAATTTACAGTTAGGAGCTAAAGCCCCAGCTAAT 2337  
Db 1136 LysAsnGluAlaLysValProLeuMetSerPheValAlaSerGlyAspGluAlaSerAla 1155  
QY 2338 GCCCATCAGATTGACTCTAGGAAT----- 2367  
Db 1156 GluIleSerAsnLeuSerValSerAspLeuGlnIleHisValValThrProGluIleGlu 1175  
QY 2368 ATGCCTAAGTATGGCTATCAAGAAAGCTGGAAGCTTGGCTGGGATCTTAATACAGCAAT 2427  
Db 1176 GluAspThrTyrGlyHisMetGlyAspTrpSerGluAla-----LysIleGln 1191  
QY 2428 AATGGTCTTACTCTGAAAGCTACATGAGCTAAAGCTGGGTATTAATCTCGGGCTGAG 2487

Db 1192 AspGly-----ThrLeuValIleSerTrpAsnProThrGlyTyrArgLeuAspProGln 1209  
QY 2488 CGAGTAGCTTCTTGGTTCCAAATAGTTATGG-----GGATCCATTTTA----- 2532  
Db 1210 LysAlaGlyAlaLeuValPheAsnAlaLeuTrpGluGluGlyAlaValLeuSerAlaLeu 1229  
QY 2533 ---GATATACATCTCGCATTCAGCAATTCACCAAGTGTGGATGGCGCTCTTATTGT 2589  
Db 1230 LysAsnAlaArgPheAlaHisAsnLeuThrAlaGlnArgMetGlu---PheAspTyrSer 1248  
QY 2590 CGAGGATTATGGGTTTCTGGAGTTTCGAATTTCTTCTCATCATGACCGCGATCTTAGGT 2649  
Db 1249 ThrAsnValTrpGlyPheAlaPheGlyGlyPheArgThrLeuSerAlaGluAsnLeuVal 1268  
QY 2650 -----CAGGATATCGGTATATTAGTGGGGTTTATCTCTTAGGAGCAAAC----- 2694  
Db 1269 AlaIleAspGlyTyrLysGlyAlaTyrGlyAlaSerAlaGlyValAspIleGlnLeu 1288  
QY 2695 ---TCCTACTTTGGATCATCATGATGTTGGTCTAGCATTTACGAAGTATTTGGTAGACT 2751  
Db 1289 MetGluAspPheValLeuGlyValSerGlyAlaAlaPheLeuGlyLysMetAspSerGln 1308  
QY 2752 AAA---GATTATGTAGTGTGTCGTTCCAAATCATCATGCTTCATAGGATCGTTTATCTA 2808  
Db 1309 LysPheAspAlaGluValSerArgLys-----GlyValValGlySerValTyrThr 1325  
QY 2809 TCTACCCCAACAGCTTTATGTGATCTCTATTGTTGGAGATGCGTTTATCCGTGCTAGC 2868  
Db 1326 -----GlyPheLeuAlaGlySerTrpPhePheLysGlyGln 1337  
QY 2869 TAGGGTTTGG-----AATCAGCATATGAAAACCTCATATATACATTTTGCAGAGAGAGC 2922  
Db 1338 TyrSerLeuGlyGluThrGlnAsnAspMetLysThrArgTyrGlyValLeuGlyLysSer 1357  
QY 2923 GATGTTCTGGGATAATAACTGCTGCTGCGTAGAGATTTGGAGCGGATTTACCGATTGTG 2982  
Db 1358 SerAlaSerTrpThrSerArgGlyValLeuAlaSerAlaLeuValGluTyrArgSerLeu 1377  
QY 2983 ATTACTCCATCTAAG-----CTCTATTGATGAGTTGCGTCTCTTTCGTCGAAGCTGAG 3036  
Db 1378 ValGlyProValArgProThrPheTyrAlaLeuHisPheAsnProTyrValGluValSer 1397  
QY 3037 TTTTCTTATGCGCATCATGAATCTTTTACAGAGAAGCGATCAAGCTCGGCGATTCAAG 3096  
Db 1398 TyrAlaSerMetLysPheProGlyPheThrGluGlnGlyArgGluAlaArgSerPheGlu 1417  
QY 3097 AGCGGACATCTCTAAATCTATCAGTCTCTGTTGGAGTGAAGTTTGTATGATCGATGT----- 3150  
Db 1418 AspAlaSerLeuThrAsnIleThrIleProLeuGlyMetLysPheGluLeuAlaPheIle 1437  
QY 3151 TCTAGTACACATCTCTAAATAATATAGCTTTTATGGCGCTTATATCTGTGATGCTTATCGC 3210  
Db 1438 LysGlyGlnPheSerGluValAsnSerLeuGlyIleSerTyrAlaTrpGluAlaTyrArg 1457  
QY 3211 ACCATCTCTGTGACTGAGACAAACGCTCTCTATCCCATCAAGAGACATGAGCAACAGATGCC 3270  
Db 1458 LysValGluGlyGlyAlaValGlnLeuLeuGluAlaGlyPheAspTrpGluGlyAlaPro 1477  
QY 3271 TTTCTATTAGCAAGCATGGAGTTGTGGTT-----AGAGATCTATG 3312  
Db 1478 MetAspLeuProArgGlnGluLeuArgValAlaLeuGluAsnAsnThrGluTrpSerSer 1497  
QY 3313 TATGCTTCTCTAAACAAGTAATATAGATATATGCGCATGGAAGATATGAGTATCGAGAT 3372  
Db 1498 TyrPheSerThrValLeuGlyLeuThrAlaPhe---CysGlyGlyPheThrSerThrAsp 1516  
QY 3373 GCTTCTCGAGGCTATGGTTGATGTCGAGGAAGTATAGTCCGGTTC 3417  
Db 1517 SerLysLeuGlyTyrGluAlaAsnThrGlyLeuArgLeuIlePhe 1531  
RESULT 10  
US-11-018-868-18

; Sequence 18, Application US/11018868  
; Publication No. US2006003487A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; APPLICANT: Bonci, Alessandro  
; APPLICANT: Finco, Oretta  
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis  
; FILE REFERENCE: 002441.00099 (PP23152.001)  
; CURRENT APPLICATION NUMBER: US/11/018,868  
; CURRENT FILING DATE: 2004-12-22  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 1531  
; TYPE: PRN  
; ORGANISM: Chlamydia trachomatis  
US-11-018-868-18

Alignment Scores:  
Pred. No.: 2,52e-43 Length: 1531  
Score: 597.50 Matches: 293  
Percent Similarity: 36.7% Conservative: 153  
Best Local Similarity: 24.1% Mismatches: 474  
Query Match: 7.6% Indels: 295  
DB: 11 Gaps: 54

US-10-701-844-1 (1-4435) x US-11-018-868-18 (1-1531)

QY	430	TCTTCTGCTCTTAAATGGGGGGATATGACGACGAATCATGTT-----	477
DB	393	SerThrAlaCysLeuGlyGlyAlaIleAlaGlnGluIleValSerIleGlnAsn	412
QY	478	---CCTCAAGGAATT---TACATGGGAGACGTTAACTGATCATTTCCCTATACTGTT	531
DB	413	AsnGlnAlaGlyIleSerPheGluGlyGlyAlaSerPheGlyGly-----	428
QY	532	ATAGGAGATCCGAGTGGGACTACTGTTTTCTCGAGGAGAGTTACATTAATAATCTT	591
DB	429	---GlyIleAlaCysGlySerPheSerAlaGlyGlyAlaSerValLeuGlyThrIle	447
QY	592	GAC-----AATCTATTGCGCTTTGCCCTTTAAGT-----TGTTTGGGAACCTTA	636
DB	448	AspIleSerIleAsnLeuGlyAlaIleSerPheSerArgThrLeuCysThrSerAsp	467
QY	637	TTAGGAGATTTTACTGTTTAGGAGGAGGACACTCGTTGACTTTTCGAGACATACGAGCT	696
DB	468	LeuGlyGlnMetGluThrGlnGly---GlyGlyAlaLeuPheGlyGluAsnIleSerLeu	486
QY	697	TCTACAAATGGGCGACCTTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGT	756
DB	487	SerGluAsnAlaGlyValLeuThr-----	494
QY	757	TTTAAGAA-----TTATCTTTTCCAATTGCAATTCATTACTTGCCTGCTGCTCT	807
DB	495	PheIleAspAsnIleValIleThrPheAlaSerAsnGlyIleLeuGlyGlyAla	514
QY	808	GCTGCAACGACTAATAGGGTAGCCAGACTCCGACGACACATCTACACCGTCTAATGGT	867
DB	515	IleLeuAlaThrGlyIleValGluIleThrAsnAsnSerGluGlyIleSerPheThrGly	534
QY	868	ACTATTATTCTAAACAGACTCTTTTGTATTACTCAATAATGAGAGTTCTCATTTCTATAGT	927
DB	535	Asn-----AlaArgAlaProGlnAlaLeuProThrGlnGluPheProLeuPheSer	552
QY	928	AAT-----TTAGTCTCTGGA-----CATGGGAGCTATAGATGCTTAAG	966
DB	553	LysLeuGluGlyArgProLeuSerSerGlyIleThrSerGlyGlyAlaIleLeuGlyArg	572
QY	967	AGCTTACGGTTCAAGGAATTAGCAAGCTTTGTGTTCTTCCAGAAATACTGCTCAAGCT	1026
DB	573	GluValAlaIleLeuHis---AsnAlaAlaValPheGluGlnAsnArgLeuGlnCys	591

QY	1027	GAT-----GGGGAGCTTGTCAAGTAGTACC	1053
DB	592	SerGluGluGluAlaThrLeuLeuGlyCysCysGlyGlyAlaValHisGlyMetAsp	611
QY	1054	AGTTTCTCTGTATGGCTAAGAGGCTCCTATTCCTTTGTAGCGAATGTTGCA-----	1107
DB	612	SerThrSerIleValGlyAsnSerSer---ValArgPheGlyAsnAsnTyrAlaMetGly	630
QY	1108	---GGAGTAAGGGGGAGGATTGCTGCTGTTTCAGGATGGCAGCAGGAGTGTCTATCA	1164
DB	631	GlnGlyValSerGlyAlaLeuLeuSer-----	640
QY	1165	TCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATPACTGCGTAGAGTTTGTAT	1224
DB	641	-----LysThrValGlnLeuAlaGlyAsnGlySerValAspPheSer	654
QY	1225	GGGAACGATAGCCGAGTAGGAGGAGGATTAC-----TCCTAGGAGACCTGCTTTC	1278
DB	655	ArgAsnIleAlaSerLeuGlyGlyAlaLeuGlnAlaSerGluGlyAsnCysGluLeu	674
QY	1279	CTGAATAATGGAACCTTG-----	1299
DB	675	ValAspAsnGlyTyrValLeuPheArgAspAsnArgValGlyValGlyAlaIle	694
QY	1300	-----TTTCTCAAC	1308
DB	695	SerCysLeuArgGlyAspValIleSerGlyAsnLysGlyArgValGluGlnSerPhe	714
QY	1309	AATGTTGCTTCTCTGTTTACATTGCTGCTAAG-----CAACCA	1347
DB	715	AsnIleAlaThrArgLeuTyrValGluGluThrValGluLysValGluGluValGluPro	734
QY	1348	ACAAGTGGACAGCTTCTAATACGAGTAATAATTAC-----	1383
DB	735	AlaProGluGlnLysAspAsnAsnGluLeuSerPheLeuGlyArgAlaGluGlnSerPhe	754
QY	1384	-----GGAGAT-----	1389
DB	755	IleThrAlaAlaAsnGlnAlaLeuPheAlaSerGluAspGlyAspLeuSerProGluSer	774
QY	1390	-----GGAGGACTATCTCTCTGT	1407
DB	775	SerIleSerSerGluGluLeuAlaLysArgArgGluCysAlaGlyGlyAlaIlePheAla	794
QY	1408	AG-----AATGCTGCGCAAGCAGGATCCCAATAC-----	1437
DB	795	LysArgValArgIleValAspAsnGlnGluAlaValPheSerAsnAsnPheSerAsp	814
QY	1438	-----TCTGATCAGTT-----TCCTTTGATGAGAG	1464
DB	815	IleTyrGlyGlyAlaIlePheThrGlySerLeuArgGluGluAspLysLeuAspGlyGln	834
QY	1465	-----GGAGTAGTTTCTTTAGTAGCAATGTAGCT	1494
DB	835	IleProGluValLeuIleSerGlyAsnAlaGlyAspValValPheSerGlyAsnSerSer	854
QY	1495	GCTGGGAAA-----GGGGAGCTATTTATGCCAAAAGCTCTCG	1533
DB	855	LysArgAspGluHisLeuProHisThrGlyGlyAlaIleCysThrGlnAsnLeuThr	874
QY	1534	GTTCCT---AATGCTGCGCTGTACAAATTTTAAAGGAATATCGCTAATGCTGGAGCG	1590
DB	875	IleSerGlnAsnThrGlyAsnValLeuPheTyrAsnAsnValAlaCysSerGlyGlyAla	894
QY	1591	ATTATTATTAGGAATCTCGAGAGCTCAGTTTATCTGCTGATTATGAGATATATTATTC	1650
DB	895	ValArgIleGluAspHisGlyAsnValLeuLeuGluAlaPheGlyGlyAspIleValPhe	914
QY	1651	GATGGGAATCTTAAAGAACACAGCAAGAAATGCTGCCGATGTTAATGCGCTAACTGTG	1710
DB	915	LysGlyAsnSerSerPheArgAlaGln-----	923



```

; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 975
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-47

Alignment Scores:
Pred. No.: 2,356-39 Length: 975
Score: 551.00 Matches: 279
Percent Similarity: 38.1% Conservative: 148
Best Local Similarity: 24.9% Mismatches: 405
Query Match: 7.0% Indels: 290
DB: 11 Gaps: 57

US-10-701-844-1 (1-4435) x US-11-103-957-47 (1-975)
Qy 388 ACCTCTTCCATAGCTCTTCTTCAATGATTCTAGCTTATTCTTGCTGCTTTAAAT 447
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 32 ThrLeuIleProLysPheLeuLeuGlyAlaLeuIleValTyrAlaProTyrSerPhe 50
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 448 GGGGGGGATATCGACGAGAATCATGGTT 492
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 51 51
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 493 GATGGGAGACGTTAACTGTATCATCTCCCTATCTGTATAGGAGATCCGAGTGGACT 552
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 65 AspArgAspThrPheThrMet 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 553 ACTGTTTTTTCGACGAGGATTAACATTAAA 606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 80 AsnTyrIleIleAsnArgLysLeuLeuSerAspPheSerLeuLeuAsnLysValSer 99
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 607 GCTTTGCTTTAAAGTTGTTGGGAACATTATTAGGGAGTTTACTGTTTAGGGAGAGA 666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 100 Ser 117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 667 CAC 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 118 SerSerAlaSerIleHisPheLysHisIleAsnIleAsnGlyPheGlyAlaGlyValPhe 137
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 721 AATACGCTGCTGAGGACTGTTTACTATTAGGGTTTAAAGAATTATCTTTTCCAT 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 138 SerGluSerSer 146
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 781 TGCAATTCAATTACTTGCCTGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCG 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 147 LeuArgLysLeuValAlaPhe 157
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 841 ACGCAACATCTACACCGTCTAATGGTACTATTATTTCTAAACAGACTCTTTGTTACTC 900
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 158 158
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 901 AATAATGAGAAGTTCTATCTATAGTAATAGTCTCTCGAGATGGGGAGCTATAGAT 960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 172 AsnAsnHisIleAlaPheArgAsnAsnIleThrLysGlyAsnGlyValIleGln 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 961 GCTAAGAGCTTAAACGGTTCAAGGAATTACGAG 993
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 192 192
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 994 1050
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 207 GlyAlaIleIlePheThrAsnAsn 217
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1051 ACCAGTTCTCTGCTATGCTAACGAGGCTCTTATTCCTTTGCTAGCGAATGTTGAGGA 1110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 218 ThrSer 226
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1111 GTAAGGGGGGGAGGTTCTGCTGTTTCAGGATGGGCAGCAGGAGTGTCTATCTACT 1170
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY 2119 TTAGCAACAATGAGTTAGCAATCTCTACCAATCTCCAGCGCAAGATTCATCCT 2178  
Db |||||  
QY 2179 GCAGTCATTGGTACACACTGCTGGTCTGTGTACAAATTAGTGGGCTT 2232  
Db :|||  
QY 2233 TTTGAGGATTTGGATGATACAGCTT 2286  
Db :|||  
QY 2287 CAAAATCAATGCTCGAATATACAGTTAGGGAGTAAGCCCGCAGCTAATGCCCCATCA 2346  
Db :|||  
QY 2347 GATTGTGACTAGGGAATGAGATCCCTAAGTATGCTATCAGGAAGCTGGAAGCTTGGC 2406  
Db :|||  
QY 2407 TGGGATCCTAATACAGCAATAATGCTCTAT 2463  
Db :|||  
QY 2464 ACTGGG---TATAATCTCTGGGCTGAGCGAGTAGCTTCTTTGGTTCCAAATAGTTATGG 2520  
Db :|||  
QY 2521 GGATCCATTTAGATATACGATCTGGC 2562  
Db :|||  
QY 2563 GCAAGTGGATGGGCGCTCTATTTGTCGAGGATTTGGTTCTGGAGTT 2613  
Db :|||  
QY 2614 ---TCGAATTTCTCTATCATGACCGCGATGCTTTAGGTCAG 2652  
Db :|||  
QY 2653 GGATATCGGTATATTTAGTGGGGTTATCTCTAGGAGCAAACTCTCTATTGGATCATCG 2712  
Db :|||  
QY 2713 ATGTTTGGCTAGCATTTACCGAAGTATTTGGTAGATCAAGATTAGTAGTGTGCTG 2772  
Db :|||  
QY 2773 TCCAATCATGCTTGCATAGGATCCGTTTAT 2805  
Db :|||  
QY 2806 ---CTATCTACCAACAAGCTTTAGTGGATCCTATTGTTGGGAGATGCGTTATCCGT 2862  
Db :|||  
QY 2863 GCTAGTACGGGTTGGGAATCAGATATGAACCTCATATACATTTGCGAGGAGAGC 2922  
Db :|||  
QY 2923 GATGTTCTGGGATATAACTGTCTGGCTGGAGAGATTGGAGCGGATTAACCGATTGTG 2982  
Db :|||  
QY 2983 ATTACTCCATTAAGCTCTATTGTAATGAGTTGGTCTCTTTCGTCGAAGCTGAGTTTCT 3042  
Db :|||  
QY 3043 TATGCGGATCATGA 3081  
Db :|||  
QY 3082 GCTCGGGCATTTCAAGAGCGACATCTCTAAATCTATCAGTTCTCTGTGGAGTGAACCTT 3141  
Db :|||  
QY 3142 GATCGATTTCTAGTACACATCTCTAATAATAGTCTTTATGCGG 3189  
Db :|||  
QY 3189 GATCGATTTCTAGTACACATCTCTAATAATAGTCTTTATGCGG 3249  
Db :|||  
QY 3249 TATATCTGTGATGCTTATCGCACCATCTCTGTGTAAGTACGAGCAAGCTCTATCCATCAA 3291  
Db :|||  
QY 3291 GAGACATGACACAGAT 3351  
Db :|||  
QY 3351 GTTGTGGTATGAGATCTATGATGCTCTCTCAAGTAAATAGATATATAGGCGCAT 3391  
Db :|||  
QY 3391 TATATCTGTGATGCTTATCGCACCATCTCTGTGTAAGTACGAGCAAGCTCTATCCATCAA 3411  
Db :|||  
QY 3411 GAGACATGACACAGAT 3471  
Db :|||  
QY 3471 GATCGATTTCTAGTACACATCTCTAATAATAGTCTTTATGCGG 3511  
Db :|||

## RESULT 12

US-11-018-868-15  
; Sequence 15, Application US/11018868  
; Publication No. US20060034871A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; APPLICANT: Bonci, Alessandro  
; APPLICANT: Finco, Oretta  
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis  
; FILE REFERENCE: 002441.00099 (pp23152.001)  
; CURRENT APPLICATION NUMBER: US/11/018,868  
; CURRENT FILING DATE: 2004-12-22  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 15  
; LENGTH: 975  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-11-018-868-15

Alignment Scores:  
Pred. No.: 2,35e-39 Length: 975  
Score: 551.00 Matches: 279  
Percent Similarity: 38.1% Conservative: 148  
Best Local Similarity: 24.9% Mismatches: 405  
Query Match: 7.0% Indels: 290  
DB: 11 Gaps: 57

US-10-701-844-1 (1-4435) x US-11-018-868-15 (1-975)

QY 388 AGCTTTTCCATAGTCTCTTTCTTCAATGATCTAGCTTATCTTGCTGCTTTAAAT 447  
Db :|||  
QY 32 ThrLeuileProLysPheLeuLeuGlyAlaLeuileValTyAlaProTySerPhe 50  
Db :|||  
QY 448 GGGGGGATATGACGACAGAAATCATGTT 492  
Db :|||  
QY 51 -----AlaGluMetGluLeuAlaileSerGlyHisLysGlnGly 64  
Db :|||  
QY 493 GATGGGAGAGCTTAACCTGATCATTTCCCTATATCTGTTATAGGAGATCCGAGTGGAGT 552  
Db :|||  
QY 65 AspArgAspThrPheThrMet 79  
Db :|||  
QY 553 ACTGTTTCTGCGAGGAGATTAACATTAATAA 606  
Db :|||

Db	80	AsnTyrIleIleAsnArgLysLeuIleLeuSerAspPheSerLeuLeuAsnLysValSer	99
Qy	607	GCTTTGGCTTTAAGTTGTTTGGGAAGCTATTATAGGAGTCTTTTACTGTTTATAGGAGAGGA	666
Db	100	Ser-----GlyGlyAlaPheArgAsnLeuAlaGlyLysIleSerPheLeuGlyLysAsn	117
Qy	667	CAC-----TCGTGTGACTTTTCGAGAACATACGACCTTCTACAAATCGGGCAGCTCTAACT	720
Db	118	SerSerAlaSerIleHisPheLysHisIleAsnIleAsnGlyPheGlyAlaGlyValPhe	137
Qy	721	AATAGCGCTGCTGATGGACTGTTTACTATTATGAGGGTTTAAAGAATATATCCCTTTCCAAAT	780
Db	138	SerGluSerSer-----11eGluPheThrAsp	146
Qy	781	TGCAAATCATTAATCTTTCGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCG	840
Db	147	LeuArgLysLeuValAlaPhe-----GlySerGluSer---	157
Qy	841	ACGACAAATCTACACCGTCTAATGGTACTATTATTCTAAACAGACTCTTTTGTGTACTC	900
Db	158	-----ThrGlyGlyIlePheThrAlaLysGluAspIleSerPheLys	171
Qy	901	AATAATGAGAAGTTCTCATTTATAGTAATATTAGTCTCTCGAGATGGGGAGCTATAGAT	960
Db	172	AsnAsnHisIleAlaPheArgAsnAsnIleThrLysGlyAsnGlyGlyValIleGln	191
Qy	961	GCTAAGAGCTTAACGGTTCAAGNATTAGCAAG-----	993
Db	192	-----LeuGlnGlyAspMetLysGlySerValSerPheValAspGlnArg	206
Qy	994	---CTTTGTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGAGCTTCTCAAGTAGTC	1050
Db	207	GlyAlaIleIlePheThrAsnAsn-----GlnAlaVal	217
Qy	1051	ACCAGTTTCTCTGCTATGCTTACACGAGGCTCTATTGCCCTTGTAGCGAATGTTCAGGA	1110
Db	218	ThrSer---SerSerMetLysHisSerGly-----	226
Qy	1111	GTAGAGGGGGAGGATGTGCTGCTTCAGGATGGGAGGAGGTGTCATCATCTACT	1170
Db	227	---ArgGlyGlyAlaIleSerGlyAspPheAlaGlySerArg-----	239
Qy	1171	TCAACAGAAGATCCAGTAGTAGTAAATTTTCCAGAAATACTCGCGTAGAGTTTGATGGGAAC	1230
Db	240	-----IleLeuPheLeuAsnAsnGlnIleThrPheGluGlyAsn	253
Qy	1231	GTAGCCGAGTAGGAGGAGGATTTACTCTTCACGGGAACGTTGCTTCTCGAATAATGGA	1290
Db	254	SerAlaValHisGlyGlyAlaIleTyrAsn-----LysAsnGly	266
Qy	1291	AAACCTTGTTCTCAACAATGTGCTCTCTCTGTTACATTTGCTGCTAGCAACCAACA	1350
Db	267	LeuValGluPheLeuGlyAsn---AlaGlyPro-----LeuAlaPheLysGluAsnThr	283
Qy	1351	-----AGTGGACAGCGCTCTTAATACGAGTAGT-----AATTACGGAGATGGA	1392
Db	284	ThrIleAlaAsnGlyGlyAlaIleTyrThrSerAsnPheLysAlaAsnGlnThrSer	303
Qy	1393	GGAGCTATCTTCTGTAAGAATGGTGGCGAAGCAGATCCAAATACTCTGGAATCATGTTCC	1452
Db	304	ProIleLeuPheSerGlnAsnHisAla-----AsnLysLysGlyGlyAlaIle	319
Qy	1453	TTT-----GATGGAGGGGAGTAGTCTTTTCTTTTCTAGTAGCAAGTGA	1491
Db	320	TyrAlaGlnTyrValAsnLeuGluGlnAsnGlnAspThrIleArgPheGluLysAsnThr	339
Qy	1492	GCTGCTGGGAAGGGGAGCTATTTATGCAAAAAGCTCTCGGTGCTTAAGTGTGGCCCT	1551
Db	340	AlaLysGluGlyGlyGlyAlaIleThrSerSerGlnCysSerIleThrAlaHisAsnThr	359
Qy	1552	GTACAATTTTAAAGAAATATCGCTAATGAT-----GGTGGAGCGCATTTATTATTA---GGA	1602



```
QY 2563 GCAAGTGTGGANGGCGCTCTATTATGTCGAGATTAATGGTTCTCGAGATT----- 2613
Db      |||||
699 ArgSerLeuIleProThrSerTyrPhe---GlyValLeuIleGlyGlyThrGlyAlaGlu 717
2614 -----TCGAAATTCCTCATCATGACCGCGATGCTTTAGGTCAG 2652
718 MetSerThrHisSerSerGluGluSerPheIleSerArgLeuGlyAlaThrGlyThr 737
2653 GGAATATCGGTATATTAGTGGGGTATTCTTAGGAGCAAACTCCTATTGGTATCG 2712
Db      |||||
738 SerIleIleArgLeuThrProSerLeuThrLeuSerGlyGly-----GlySerHis 754
2713 ATGTTTGGCTAGCATTTACCAAGATATTGGTAGATCTAAAGATTAATGTAGTGTGCT 2772
Db      |||||
755 MetPheGlyAspSerPhe-----ValAlaAspLeu 764
2773 TCCAAATCATCATGCTGTGATAGGATCCGTTTAT----- 2805
765 ProGluHisIleThrSerGluGlyIleValGlnAenValGlyLeuThrHisValTrpGly 784
2806 ---CTATCTACCAACAAGCTTTATGTGATCTCTATTGTTGGAGATGCGTTTATCGGT 2862
Db      |||||
785 ProIleuThrValAenSerThrLeuCysAlaIle---LeuAspHisAenAlaMetValArg 803
2863 GCTAGTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAGAGC 2922
Db      |||||
804 IleCys-----SerIlySlyAspHisThrTyrGly----- 813
2923 GATGTTCTGGGATAATAACTGCTGCTGGAGAGATGAGCGGGATTAACGATTGTG 2982
Db      |||||
814 ---LysTrpAspThrPheGlyMetArgGlyThrLeuGlyAlaSerTyrThrPheLeu 831
2983 ATTACTCCATCTAAGCTCTATTGAATGAGTTGGCTCTCTTCGTCGAAGCTGAGTTTCT 3042
Db      |||||
832 -----GluTyrAspGlnThrMetArg-----ValPheSer 841
3043 TATGCCGATCATGAA-----TCTTTTACAGAGAGGCGATCAA 3081
Db      |||||
842 PheAlaAenIleGluAlaThrAenIleLeuGlnArgAlaPheThrGluThrGlyTyrAen 861
3082 GCTGGGATTTCAAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGAGTGAAGTTT 3141
Db      |||||
862 ProArgSerPheSerLysThrIlySleLeuAenIleAlaIleProIleGlyIleGlyTyr 881
3142 GATCATGTTCTAGTACACATCTCTATAAATATAGCTTTATGGCG-----GCT 3189
Db      |||||
882 GluPheCysLeuGlyAen-----SerSerPheAlaLeuLeuGlyIlyGlySerIleGly 899
3190 TATATCTGTGATGCTTATGCGACCATCTCTGGTACTGAGACAACTCTATCCCATCAA 3249
Db      |||||
900 TyrSerArgAspIleIlySArgGluAenProSerThrLeuAlaHisLeuAlaMetAenAsp 919
3250 GAGACATGAGACACAGAT-----GCCCTTTCATTTAGCAAGACATGA 3291
Db      |||||
920 PheAlaTrpThrThrAenGlyCysSerValProThrSerAlaHisThrLeuAlaAenGln 939
3292 GTTGGTGTAGGATCTATGATCTCTCTCTAAAGTAATATAGAGTATATGCCAT 3351
Db      |||||
940 LeuIleLeuAen-----TyrIlySAlaCysSerLeuTyrIleThrAlaTyrThrIle 956
3352 GGAAGATATGATATCGAGATCTCTCGAGGCTATGTTGAGTGCAGGAGTAGTAGTC 3411
Db      |||||
957 AenArg---GluGlyIlyAenLeuSerAen-----SerLeuSerCysGlyGlyTyrVal 973
3412 CGGTTTC 3417
974 GlyPhe 975
```

RESULT 13

US-11-018-868-50

; Sequence 50, Application US/11018868

; Publication NO. US20060034871A1

```
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (P23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50
; LENGTH: 975
; TYPE: PRN
; ORGANISM: Chlamydia trachomatis
US-11-018-868-50
```

```
Alignment Scores:
Pred. No.: 2,35e-39 Length: 975
Score: 551.00 Matches: 279
Percent Similarity: 38.1% Conservative: 148
Best Local Similarity: 24.9% Mismatches: 405
Query Match: 7.0% Indels: 290
DB: 11 Gaps: 57
```

US-10-701-844-1 (1-4435) x US-11-018-868-50 (1-975)

```
QY 388 ACGTCTTCCATAAGTTCTTTTCAATGATTCTAGCTTATTCTTGCTGCTCTTTAAAT 447
Db      |||||
32 ThrLeuIleProLysPheLeuLeuGlyAlaLeuIleValTyrAlaProTyrSerPhe--- 50
448 GGGGGGGATATGAGCAGCAAAATCATGTT-----CCTCAAGGAATTTTAC 492
Db      |||||
51 -----AlaGluMetGluLeuAlaIleSerGlyHisLysGlnGly---Lys 64
493 GATGGGAGAGCTTAACGTGTATCATTTCCCTATCTATGTTATAGGAGATCCGAGTGGACT 552
Db      |||||
65 AspArgAspThrPheThrMet-----IleSerSerCysProGluGlyThr 79
553 ACTGTTTCTTTCGCGAGGAGGTAAACATTAAAA-----AATCTTGACAAATTTCTATTCGA 606
Db      |||||
80 AsnTyrIleIleAenArgLysLeuIleLeuSerAspPheSerLeuLeuAenLysValSer 99
607 GCTTTGCTTTAAGTTGTTTGGGAACCTATTAGGAGTCTTTACTGTTTATAGGGAGAGGA 666
Db      |||||
100 Ser-----GlyGlyAlaPheArgAenLeuAlaGlyLysIleSerPheLeuGlyLysAen 117
667 CAC-----TCGTTGACCTTTCGAGAACATACGGACTTCTACAAATGGGCGAGCTTAAGT 720
Db      |||||
118 SerSerAlaSerIleHisPheLysHisIleAenIleAenGlyPheGlyAlaGlyValPhe 137
721 AATAGCGCTGCTGATGGACTGTTTACTATTGAGGGTTTTTAAAGAATTATCTTTTCCCAAT 780
Db      |||||
138 SerGluSerSer-----IleGluPheThrAsp 146
781 TGCAATTCATTACTTTCGCGTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCG 840
Db      |||||
147 LeuArgLysLeuValAlaPhe-----GlySerGluSer--- 157
841 ACGCAACATCTACACCGCTTAATGGTACTATTATTCTTAAACAGATCTTTTGTACTCTC 900
Db      |||||
158 -----ThrGlyGlyIlePheThrAlaLysGluAspIleSerPheLys 171
901 AATATGAGAAGTTCTCAATCTATAGTAATTTAGTCTCTGAGAGATGGGGAGCTATAGAT 960
Db      |||||
172 AenAenHisHisIleAlaPheArgAenAenIleThrLysGlyAenGlyValIleGln 191
961 GCTAAGAGCTTAAACGGTTTCAAGGAATTTAGCAAG----- 993
Db      |||||
192 -----LeuGlnGlyAspMetLysGlySerValSerPheValAspGlnArg 206
994 ---CTTTGCTCTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTC 1050
```



2983	QY	ATTACTCCATCTAAGCTCTATTGAATGAGTTCGGTCCCTTCGTGCAAGCTGAGTTTTCT	3042
832	Db	 :::     -----GluTyrAspGlnThrMetArg-----ValPheSer	841
3043	QY	TATGCCGATCATGAA-----TCCTTTTACAGAGGAAGCGCATCAA	3081
842	Db	:::    :      PheAlaAsnIleGluAlaThrAsnIleLeuGlnArgAlaPheThrGluThrGlyTyrAsn	861
3082	QY	GCTCGGCATCTCAAGAGCGGCACATCTCCTAAATCTATCAGTTCCTGTTGAGGAGAGTTT	3141
862	Db	 :::    : -----ProArgSerPheSerLysThrLysLeuLeuAsnIleAlaIleProIleGlyIleGlyTyr	881
3142	QY	GATCGATGTTCTAGTACACATCCCTAAATAATATAGCTTTATGGCG-----GCT	3189
882	Db	:::     :::    : -----SerSerPheAlaLeuLeuGlyLysGlySerIleGly	899
3190	QY	TATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACAACGCTTCATCCCATCAA	3249
900	Db	 :::     TyrSerArgAspIleLysArgGluAsnProSerThrLeuAlaHisLeuAlaMetAsnAsp	919
3250	QY	GAGACATGGACAACAGAT-----GCCTTTCATTAGCAAGACATGGA	3291
920	Db	 :::    : PheAlaIrrThrThrAsnGlyCysSerValProThrSerAlaHisThrLeuAlaAsnGln	939
3292	QY	GTTCTGGTTAGAGGATCTATGATGCTTCTCTCAACAAGTAATAPAGAAGTATATGGCCAT	3351
940	Db	:::    :      -----TyrLysAlaCysSerLeuTyrlleThrAlaTyrThrIle	956
3352	QY	GGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGGTCCGAGGAAGTAGATC	3411
957	Db	 :::    : AsnArg--GluGlyLysAsnLeuSerAsn-----SerLeuSerCysGlyGlyTyrVal	973
3412	QY	CGGTTTC	3417
974	Db	 GlyPhe	975

```

RESULT 14
US-11-103-957-11
, Sequence 11, Application US/11103957
, Publication No. US20050281847A1
, GENERAL INFORMATION:
, APPLICANT: Barthet, Francois-Xavier Jacques
, APPLICANT: Lobet, Yves
, APPLICANT: Poolman, Jan
, APPLICANT: Veriant, Vincent Georges Christian Louis
, TITLE OF INVENTION: Vaccine Composition
, FILE REFERENCE: B45261
, CURRENT APPLICATION NUMBER: US/11/103,957
, CURRENT FILING DATE: 2005-04-12
, PRIOR APPLICATION NUMBER: US/10/467,534
, PRIOR FILING DATE: 2004-02-03
, PRIOR APPLICATION NUMBER: PCT/EP02/01356
, PRIOR FILING DATE: 2002-02-08
, PRIOR APPLICATION NUMBER: GB 0103169.9
, PRIOR FILING DATE: 2001-02-08
, NUMBER OF SEQ ID NOS: 108
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 11
, LENGTH: 1034
, TYPE: PR1
, ORGANISM: Chlamydia trachomatis
US-11-103-957-11

```

US-10-701-844-1 (1-4435) x US-11-103-957-11 (1-1034)

QY	421	CTAGCTATTACTTCTGCTGCTCTTTAAATGGGGGGGATATGCACGACAAATCATGCTTCCT	480
DB	7	LeuSerPheAlaCysLeuSerPhe-----PheTyLeuSerThrIleSerIleLeu	23
QY	481	CAAGGAATTTACGATGGGAGACGTTAACTGTATCATTT-----	519
DB	24	GlnAla-----AenGluThrAspThrLeuGlnPheArgPheThrPheSerAsp	40
QY	520	---CCCTATACTGTTATAGAGACTCCGAGTGGGACTACTGTTTTTCTCGACGAGAGTTA	576
DB	41	ArgGluIleGlnPheValLeuAspProAla-----SerLeuIleThrAlaGlnAsnIle	58
QY	577	ACATTAAAAAATCTGACAAATCTATTGCGACTTTGCCTTTAAAGTTGTTTGGGAACCTTA	636
DB	59	ValLeuSerAsnLeuGlnSer-----AsnGly	67
QY	637	TTAGGGAGTTTTACTGTTTTTAGGGAGGAGCACTCGTTGACTTTTCGAGAACATACCGACT	696
DB	68	ThrGlyAlaCysThrIleSerGlyAsnThrGlnIlePheSerAsn-----	84
QY	697	TCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG-----TTTACT	747
DB	85	SerValAsnThrThrAlaAspSerGlyAlaPheAspMetValThrThrSerPheThr	104
QY	748	ATTGAGGCTTTTAAAGAAATATCTCTTTCCAAATTGCAATTCATTACTTGCCTGCTGCCT	807
DB	105	AlaSerAspAsnAlaLeuLeuPhe-----CysAsnAsnTyr-----	117
QY	808	GCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACAACTCTACACCGTCTAATGCT	867
DB	118	---CysThrHisAsnIleGly-----GlyGly	125
QY	868	ACTATTATTCTAAACAGACTCTTTGTACTCAATAATGAGAAGTTCTCATTTCTATAGT	927
DB	126	AlaIleArgSerGlyGlyProIleArgPheLeuAsnAsnGlnAspValLeuPheTyrAsn	145
QY	928	AATTAGTCTCTCGA-----GATGGG	948
DB	146	AsnIleSerAlaGlyAlaLysTyValGlyThrGlyAspHisAsnGluLysAsnAspGly	165
QY	949	GGACTATAGTGTAAAGACTTAAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAA	1008
DB	166	GlyAlaLeuTyAlaThrThrIleThrLeuThrGly---AsnArgThrLeuAlaPheIle	184
QY	1009	GAATAATCTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCCACAGTTCTCTGCT---	1065
DB	185	AsnAsnMetSerGlyAspCysGlyGlyAlaIleSerAlaAspThrGlnIleSerIleThr	204
QY	1066	-----ATGGCTAACGGGCTCTATTATGCC	1089
DB	205	AspThrValLysGlyIleLeuPheGluAsnAsnHisThrLeuAsnHisIleProTyThr	224
QY	1090	TTTCTAGCGAATGTTGACGAGTAAGAGGGGGAGGATTCGTGCTGTTCAGGATGGGCAG	1149
DB	225	GlnAlaGluAsnMetAla-----ArgGlyGlyAlaIleCysSerArgAspAsp-----	240
QY	1150	CAGGAGTGTCTATCATCTTCTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACT	1209
DB	241	---LeuCysSerIleSerAsnAsnSerGlyProIleVal-----	252
QY	1210	CGGTAGATGTTTGAATGGAACTGATAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC	1269
DB	253	-----PheAsnTyAsnGlnGlyGlyLysGlyAlaIleSerAlaThrArgCys	269
QY	1270	GTTGCTTTCTCGAATAATGGAAAACTTGTCTTCTCAACAAATGTTGCTTCTCTGTTTAC	1329
DB	270	ValIleAspAsnAsnLysGluArgIleIlePheSerAsnAsn-----	283
QY	1330	ATTGCTGTAAACCAACCAACGATGGGACAGGCTTCTTAATACGAGTAATAATACCGAGAT	1389
DB	284	-----SerSerLeuGlyTrpSerGlnSerGlnSerSerAlaSerAlaSer	296

1390 GGAGGAGTATCTTCTGTAAGAATCGTGGCGCAGCAGGATCCAAATAACTCTGGATCAGTT 1449  
Db |||||  
297 GlyGlyAlaIleGlnThrThrGlnGlyPheThrLeu---ArgAsnAsnLysGlySerIle 315  
Qy |||||  
1450 TCCCTTTGATGGAGAGGAGTACTGTTCTTTAGTAGCAATGTAGCTGCGGAAAGGGGA 1509  
Db |||||  
316 TyrPheAsp-----SerAsnThrAlaThrHisAlaGlyGly 327  
Db |||||  
1510 GCTATTATGCCAAAGAGCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGTAAT 1569  
Qy |||||  
328 AlaIleAsnCyGlyTyrrileAspIleArgAsnGlyProValTyrPheLeuAsnAen 347  
Db |||||  
1570 ATCGCTAATGATGGCGGAGTATTATTAGAGAA-----TCTGGAGAGCTCAGTTTAA 1623  
Qy |||||  
348 SerAlaAlaTrpGlyAlaAlaPheAsnLeuSerLysProArgSerAlaThrAsnTyrrile 367  
Db |||||  
1624 TCTGCTGATATTGGAGATATTATTTTC-----GATGGGAAT 1659  
Qy |||||  
368 HisThrGlyThrGlyAspIleValPheAsnAsnValValPheThrLeuAspGlyAen 387  
Db |||||  
1660 CTTAAAGAACACCCAAAGAGATGCTGCCGATGTTAAATGGCGTAACTGTGCTCCACAA 1719  
Qy |||||  
388 Leu-----LeuGlyLysArgLysLeuPheHisIleAsnAsnGlnIleThrProTyr 405  
Db |||||  
1720 GCCATTTCGATCGGATCGGAGGAAATAAACACATTAAAGAGCTAAAGCAGGCGCATCAG 1779  
Qy |||||  
406 ThrLeuSerLeuGly-----AlaLysLysAspThrArg 416  
Db |||||  
1780 ATTCTCTTTAAATGAT-----CCCATCGAGATGGCAACCGGAATAACACAG 1824  
Qy |||||  
417 IleTyrPheTyrAspLeuPheGlnTrpGluArgValLysGlnAsnThrSerAsnAsnPro 436  
Db |||||  
1825 CCAGCCAGCTTTCAAA---CTTCTAAATAATTAAGATGCTGCAAGATACACAGGGAT 1881  
Qy |||||  
437 ProSerProThrSerArgAsnThrIleThrValAsnProGluThrGluPheSerGlyAla 456  
Db |||||  
1882 ATTGTTTTTGCT---NATGGAACAGTACTTTGTACCAAAATGTT----- 1923  
Qy |||||  
457 ValValPheSerTyrAsnGlnMetSerSerAspIleArgThrLeuMetGlyLysGluHis 476  
Db |||||  
1924 -----ACGATAGACAAGGAGGATTTGTTCTTCGTGAAAG 1959  
Qy |||||  
477 AsnTyrrileLysGluAlaProThrThrLeuLysPheGlyThrLeuAlaIleGluAspAsp 496  
Db |||||  
1960 GCAAAATTAACAGTAATCTTAAGTCAGACAGGTTGGAGT-----CTGTATATG 2010  
Qy |||||  
497 AlaGluLeuGluIlePheAsnIleProPheThrGlnAsnProThrSerLeuLeuAlaLeu 516  
Db |||||  
2011 GAAGCTGGGAGTACATGGGATTTGTAACTCCACACCACACAGCCTCCTCGCGCT 2070  
Qy |||||  
517 GlySerGlyAlaThrLeuThrVal-----GlyLys 526  
Db |||||  
2071 AATCAGTTGATCAGCTTTCATCTGCAATTTGCTCTTTCTTTTGTAGCAACAAT 2130  
Qy |||||  
527 HisGlyLysLeuAsnIleThrAsnLeuGlyValIleLeuProIleIleLeuLysGluGly 546  
Db |||||  
2131 GCAGTTACGAATCTCTCT-----ACCAATCTCCAGCGCAAGATTTCTCATCTGCA 2181  
Qy |||||  
547 -----LysSerProProCysIleArgValAsnProGlnAspMetThrGlnAsnThrGly 564  
Db |||||  
2182 GTCATGGTAGCACACTGCTGTTCTGTTACAAATTTAGTGGCCCTACTCTTTT----- 2235  
Qy |||||  
565 Thr-----GlyGlnThrProSerSerThrSerSerIleSerThrProMetIleIlePheAsn 583  
Db |||||  
2236 -----GAGGATTTGATGATACAGTTATGATAGTATGATGTTGGCTA 2277  
Qy |||||  
584 GlyArgLeuSerIleValAspGluAsnTyrrGluSerValTyrAspSerMetAspLeu--- 602  
Db |||||  
2278 GGTTCATCAAAATAATCAATGCTCCTGAAATTAACAGTTAGGAGTAAAGCCCCAGCTAAT 2337  
Qy |||||  
603 ---SerArgGlyLysAlaGluGlnLeuIleLeuSerIleGluThrThrAsnAspGlyGln 621  
Db |||||  
2338 GCCCCATCAGT-----TTGACTCTAGGGAATCAGATG-----CCTAAGTATGGC 2382  
Qy |||||

622 LeuAspSerAsnTrpGlnSerSerLeuAsnThrSerLeuLeuSerProHisTyrrGly 641  
Db |||||  
2383 TATCAGCAAGCTCGAAGCTTGGCGGATCCTAATACA----- 2421  
Qy |||||  
642 TyrGlnGlyLeuTrpThrProAsnTrpIleThrThrThrIleThrLeuAsnAen 661  
Db |||||  
2421 ----- 2421  
Qy -----  
662 AsnSerSerAlaProThrSerAlaThrSerIleAlaGluGlnLysLysThrSerGluThr 681  
Db -----  
2421 ----- 2421  
Qy -----  
682 PheThrProSerAsnThrThrThrAlaSerIleProAsnIleLysAlaSerAlaGlySer 701  
Db -----  
2422 -----GCAAAATAATCGT-----CCTTATACTCTGAAAGCT 2451  
Qy -----  
702 GlySerGlySerAlaSerAsnSerGlyGluValThrIleThrLysHisThrLeuValVal 721  
Db -----  
2452 ACATCGACTAAAACTGGGTATTAATCTGGCGCTGAGCGAGTAGCTTTCTTTGGTTCCAAAT 2511  
Qy |||||  
722 AsnTrpAlaProValGlyTyrrileValAspProIleArgGlyAspLeuIleAlaAsn 741  
Db |||||  
2512 AGTTTATGGGGATCCATTTTATAGATATACGATCTGGCATTCAGCAATTCAGCAAGTGTG 2571  
Qy |||||  
742 SerLeu-----ValHisSer----- 746  
Db -----  
2572 GATGGCGCTCTTATTGTCGAGGATTA-----TGGGTTCT 2607  
Qy |||||  
747 ---GlyArgAsnMetThrMetGlyLeuArgSerLeuLeuProAsnAsnSerTrpPheAla 765  
Db -----  
2608 -----GGAGTTTTCGAATTTCTTATCATGACGCGATGCTTTAGT---CAGGGA 2655  
Qy -----  
766 LeuGlnGlyAlaAlaThrThrLeuPheThrLysGlnLysArgLeuSerTyrHisGly 785  
Db -----  
2656 TATCGTATATTAGTGGGTTATCTCTTAGGAGCAAACTCTTACTTTGATCATCATG 2715  
Qy |||||  
786 TyrSerSerAlaSerLysGlyTyrrValSerSerGlnAla-----SerGlyAla 802  
Db -----  
2716 TTTGCTCTAGCATTTACCGAAGTATTGCTAGATCTAAAGATTATGCTGCTCGTTCC 2775  
Qy |||||  
803 HisGlyHisLysPheLeuLeuSerPheSerGlnSerSerAspLysMetLysGluLysGlu 822  
Db -----  
2776 AATCATCATGCTGTCATAGATCCGTTTATCTATCTATCTACCCACAAAGCTTTTATGATCC 2835  
Qy -----  
823 ThrAsnAsnArgLeuSerSerArgTyrTyrLeuSer-----AlaLeuCysPheGlu 839  
Db -----  
2836 TATTTGTTCCGAGAT-----GCGTTTATC---CGTGCTAGTACCGGTTTGGGAATCAG 2886  
Qy -----  
840 HisProMetPheAspArgIleAlaLeuIleGlyAlaAlaCysAsnTyrGlyThrHis 859  
Db -----  
2887 CATATGAAAACCTCATATACATTTGCGAGGAGGAGGATGTTGTTGGGATAATACTGT 2946  
Qy -----  
860 AsnMetArgSerPheTyrGlyThrLysLysSerSerLysGlyLysPheHisSerThrThr 879  
Db -----  
2947 CTGCTGGAGAGATTGCGCGGATTTACCGATTGTGATTAATCTCATCT---AAGCTCTAT 3003  
Qy -----  
880 Leu-----GlyAlaSerLeuArgCysGluLeuArgAspSerMetProLeuArg 895  
Db -----  
3004 TTGAATGAGTTCCTTCTTCGCAAGCTGAGTTTCTTTATGCGGATCATGATCTTTT 3063  
Qy -----  
896 SerIleMetLeuThrProPheAlaGlnAlaLeuPheSerArgThrGluProAlaSerIle 915  
Db -----  
3064 ACAGAGAGGCGCATCAAGCTCCGGCATTC-----AAGAGCGGACATCTCTAAATCTA 3117  
Qy -----  
916 ArgGlySerGlyAspLeuAlaArgLeuPheThrLeuGluGlnAlaHisThrAlaValVal 935  
Db -----  
3118 TCAGTTCTCTGAGTGAAGTTTGTATCGATGTTCTAGTACATCATCTTAAT---AAATAT 3174  
Qy -----  
936 Ser---ProIleGlyIleLysGlyAlaTyrSerSerAspThrTrpProThrLeuSerTrp 954  
Db -----  
3175 AGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTCTGTTACTGAGACAAG 3234  
Qy -----

Db 955 GluMetGluLeuAlaTyrglnProThrLeuTyrrTpLySArgProLeuLeuAenThrLeu 974  
 Qy 3235 CTCCTATCCCATCAAGAGACATGGACAACAGATGCTTTTCATTTTAGCAAGACATGAGTT 3294  
 Db 975 LeuileGlnAenAenGlySerTpValThrThrAenThrProLeuAlaLysHisSerPhe 994  
 Qy 3295 GTGGTTAGAGATCTATGATGCTTCTCTAACAGTAATATAGAAGTATATGGCCATGA 3354  
 Db 995 TyrGlyArgGlySer---HisSerLeuLysPheSerHisLeuLysLeuPheAlaAenTyr 1013  
 Qy 3355 AGATGATGATCGAGATGCTTCTCGAGGCTATGCTTTCAGTGCAGGAGTAGATCGCG 3414  
 Db 1014 GlnAlaGluValAlaThrSerThrValSerHisTyrIleAenAlaGlyGlyAlaLeuVal 1033  
 Qy 3415 TTC 3417  
 Db 1034 Phe 1034  
 RESULT 15  
 US-11-018-868-20  
 / Sequence 20, Application US/11018868  
 / Publication No. US20060034871A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Chiron Corporation  
 / APPLICANT: Grandi, Guido  
 / APPLICANT: Ratti, Giulio  
 / APPLICANT: Bonci, Alessandro  
 / APPLICANT: Finco, Oretta  
 / TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis  
 / FILE REFERENCE: 002441.00099 (P23152.001)  
 / CURRENT APPLICATION NUMBER: US/11/018,868  
 / CURRENT FILING DATE: 2004-12-22  
 / NUMBER OF SEQ ID NOS: 186  
 / SOFTWARE: PatentIn version 3.2  
 / SEQ ID NO 20  
 / LENGTH: 1034  
 / TYPE: PR1  
 / ORGANISM: Chlamydia trachomatis  
 US-11-018-868-20

Alignment Scores:  
 Pred. No.: 7, 41e-37 Length: 1034  
 Score: 522.50 Matches: 274  
 Percent Similarity: 37.1% Conservative: 157  
 Best Local Similarity: 23.6% Mismatches: 435  
 Query Match: 6.6% Indels: 295  
 Db: 11 Gaps: 53

US-10-701-844-1 (1-4435) x US-11-018-868-20 (1-1034)

Qy 421 CTAGCTTATCTGCTGCTCTTTAAATGGGGGGATATGCAGCAGAAATCATGTTCTCT 480  
 Db 7 LeuSerPheAlaCysLeuSerPhe-----PheTyrLeuSerThrIleSerIleLeu 23  
 Qy 481 CAGGAATTTACGATGGGAGACGTTAATGATCATTT----- 519  
 Db 24 GlnAla-----AsnGluThrAspThrLeuGlnPheArgPheThrPheSerAsp 40  
 Qy 520 ---CCCTATCTGTTATAGGACATCGAGTGGGACTGTTCTTTTCTGCAGGAGTTA 576  
 Db 41 ArgGluIleGlnPheValLeuAspProAla-----SerLeuIleThrAlaGlnAenIle 58  
 Qy 577 ACATTAATAAAATCTTGACAAATCTATGTCAGCTTTGCTTCTTAAAGTTGTTGGGAACTTA 636  
 Db 59 ValLeuSerAsnLeuGlnSer-----AsnGly 67  
 Qy 637 TTAGGAGTTTCTGTTTATAGGAGAGACACTCGTTGACTTTTCGAGACATACGGACT 696  
 Db 68 ThrGlyAlaCysThrIleSerGlyAenThrGlnIlePheSerAen----- 84  
 Qy 697 TCTACAAATGGGACGCTTAAGTAATAGCGCTGCTGATGGAATG-----TTTACT 747  
 Db 85 SerValAenThrThrAlaAspSerGlyGlyAlaPheAspMetValThrThrSerPheThr 104

Qy 748 ATTGAGGGTTTAAAGAAATTATCCTTTTCCAAATTCGAATTCATTACTTGCCTACTGCCT 807  
 Db 105 AlaSerAspAenAlaAenLeuPhe-----CysAenAenTyr----- 117  
 Qy 808 GCTCGAACGACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACACCGCTTAATGGT 867  
 Db 118 ---CysThrHisAenLysGly-----GlyGly 125  
 Qy 868 ACTATTATTCTAAAAACAGATCTTTTGTACTCAATATAGAGAATTCCTATTCTATAGT 927  
 Db 126 AlaIleArgSerGlyGlyProIleArgPheLeuAenAenGlnAenValLeuPheTyrAen 145  
 Qy 928 AATTAGTCTCTGCA-----GATGGG 948  
 Db 146 AenIleSerAlaGlyAlaLysTyrValGlyThrGlyAspHisAenGluLysAenArgGly 165  
 Qy 949 GGAGCTATAGATCTTAAGAGCTTAACGCTTCAAGGAATTAGCAAGCTTTGTGCTTCCAA 1008  
 Db 166 GlyAlaLeuTyrAlaThrThrIleThrLeuThrGly---AsnArgThrLeuAlaPheIle 184  
 Qy 1009 GAAAATACTGCTCAAGCTGATGGGGAGCTTTGTCAAGTAGTACCAGTTTCTCTGCT--- 1065  
 Db 185 AsnAenMetSerGlyAspCysGlyGlyAlaIleSerAlaAspThrGlnIleSerIleThr 204  
 Qy 1066 -----ATGGCTAACGAGGCTCTTATGCT 1089  
 Db 205 AspThrValLysGlyIleLeuPheGluAenAenHisThrLeuAenHisIleProTyrThr 224  
 Qy 1090 TTTGTAGCGAATGTTGCAGGAGTAAGAGGGGAGGATGCTGCTGTTTCAGAGTGGGAG 1149  
 Db 225 GlnAlaGluAenMetAla-----ArgGlyGlyAlaIleCysSerArgArgAsp----- 240  
 Qy 1150 CAGGAGTGTCTATCTTCTCAACAGAGATCCAGTAGTAAGTTTTCAGAGAAATACT 1209  
 Db 241 -----LeuCysSerIleSerAenAenSerGlyProIleVal----- 252  
 Qy 1210 CGGTAGATGTTGATGGGAACGTAGCCGAGTAGAGGAGGAGGATTTACTCTCGGGAAC 1269  
 Db 253 -----PheAenTyrAsnGlnGlyGlyLysGlyGlyAlaIleSerAlaThrArgCys 269  
 Qy 1270 GTTGTCTTCTCGAATAATGGAACCTTTCTCAACAATGTTGCTTCTCTCTGTTTAC 1329  
 Db 270 ValIleAspAenAenLysGluArgIleIlePheSerAenAen----- 283  
 Qy 1330 ATTGCTGCTAAGCAACCAAGTAGCGACAGCTTCTAATACGAGTAATAATACGAGAT 1389  
 Db 284 -----SerSerLeuGlyTyrSerGlnSerSerSerAlaSerAen 296  
 Qy 1390 GGAGAGACTATCTCTGTGAAGATGGTGGCAGCAGGATCCAATACTCTGGATCAGTT 1449  
 Db 297 GlyGlyAlaIleGlnThrThrGlnGlyPheThrLeu---ArgAenAenLysGlySerIle 315  
 Qy 1450 TCCTTTGATGGAGGAGTAGTTTCTTTAGTAGCAATGATGCTGCTGGAAAGGGGA 1509  
 Db 316 TyrPheAsp-----SerAenThrAlaThrHisAlaGlyGly 327  
 Qy 1510 GCTATTTATGCCAAAAAGCTCTCGTGTCTAACTGTGCTCTGTACAAATTTTAAAGGAAT 1569  
 Db 328 AlaIleAenCysGlyTyrIleAspIleArgAenGlyProValTyrPheLeuAenAen 347  
 Qy 1570 ATCGCTAATGATGGTGGAGCGATTTATTAGAGAA-----TCTGGAGAGCTCAGTTTA 1623  
 Db 348 SerAlaIleTyrGlyAlaAlaPheAenLeuSerLysProArgSerAlaThrAenTyrIle 367  
 Qy 1624 TCTGCTGATTATGGAGATATTATTTC-----GATGGGAAT 1659  
 Db 368 HisThrGlyThrGlyAspIleValPheAenAenAenValValPheThrLeuAspGlyAen 387  
 Qy 1660 CTTAAAGAACACCAAGAGAGATGCTCCGATGTTAATGGCGTAACGTGTGCTCCACAA 1719  
 Db 398 Leu-----LeuGlyLysArgLysLeuPheHisIleAenAenAenGluIleThrProTyr 405

1720 GCCATTTTCAGTGGGATCGGGAGGAAATAACGACATTAAAGCTAAAGCAGGCGATCAG 1779  
Db : : : : :  
406 ThrLeuSerLeuGly-----AlaIysAspThrArg 416  
1780 ATTCTCTTAATGAT-----CCCATCGAGATGGCAAAACGGAATAAACCGAG 1824  
Db : : : : :  
417 IleTyrPheTyrAspLeuPheGlnTrpGluArgValLysGluAsnThrSerAsnAsnPro 436  
1825 CAGGCGAGCTCTCCAAA---CTTCTAAAATAACGATGGTGAAGGATACACAGGGAT 1881  
Db : : : : :  
437 ProSerProThrSerArgAsnThrIleThrValAsnProGluThrGluPheSerGlyAla 456  
1882 ATTGTTTTCGCT---AATCGAAGCAGTACTTTGTACCAAAATGCTT----- 1923  
Db : : : : :  
457 ValValPheSerTyrAsnGlnMetSerSerAspIleArgThrLeuMetGlyLysGluHis 476  
1924 -----ACGATAGAGCAAGGAGGATTTCTTCTCGTGAAGAG 1959  
Db : : : : :  
477 AsnTyrIleLysGluAlaProThrThrLeuLysPheGlyThrLeuAlaIleGluAspAsp 496  
1960 GCAAAATTATCAGTGAATCTCTAAGTCAGACAGGTGGAGT-----CTGTATATG 2010  
Db : : : : :  
497 AlaGluLeuGluIlePheAsnIleProPheThrGlnAsnProThrSerLeuLeuAlaLeu 516  
2011 GAAGCTGGGAGTACATGGGATTTGTAACTCCACACACCACAGCCTCTCGCGCT 2070  
Db : : : : :  
517 GlySerGlyAlaThrLeuThrVal-----GlyLys 526  
2071 AATCAGTTCATCAGCTTCCCAATCTGCATTTGTCTTTCTTTTGTAGCAACAAT 2130  
Db : : : : :  
527 HisGlyLysLeuAsnIleThrAsnLeuGlyValIleLeuProIleLeuLysGluGly 546  
2131 GCAGTTACGAATCTCTCT-----ACCAATCTCCAGCGGCAAGATCTCATCTCGCA 2181  
Db : : : : :  
547 -----LysSerProProCysIleArgValAsnProGlnAspMetThrGlnAsnThrGly 564  
2182 GTCATTTAGTACCACTGCTGTTCTGTACAAATTAGTGGCTATCTTTT----- 2235  
Db : : : : :  
565 Thr---GlyGlnThrProSerSerThrSerIleSerThrProMetIleIlePheAsn 583  
2236 -----GAGGATTTGATGATACAGCTTATGATAGGATGATGATGCTA 2277  
Db : : : : :  
584 GlyArgLeuSerIleValAspGluAsnTyrGluSerValTyrAspSerMetAspLeu--- 602  
2278 GGTCTTAATCAAAAATCAATGCTCGAATTAACAGTTAGGAGTAAAGCCCCAGCTAAT 2337  
Db : : : : :  
603 ---SerArgGlyLysAlaGluGlnLeuIleLeuSerIleGluThrThrAsnAspGlyGln 621  
2338 GCCCATCAGAT-----TTGACTCTAGGGAATCAGATG-----CCTAAGTATGCG 2382  
Db : : : : :  
622 LeuAspSerAsnTrpGlnSerSerLeuAsnThrSerLeuLeuSerProProHisTyrGly 641  
2383 TATCAAGAGAGCTGGAAGCTTGGCTGGATCTCTAATAACA----- 2421  
Db : : : : :  
642 TyrGlnGlyLeuTrpThrProAsnTrpIleThrThrTyrThrIleThrLeuAsnAsn 661  
2421 ----- 2421  
662 AsnSerSerAlaProThrSerAlaThrSerIleAlaGluGlnLysLysThrSerGluThr 681  
2421 ----- 2421  
682 PheThrProSerAsnThrThrThrAlaSerIleProAsnIleLysAlaSerAlaGlySer 701  
2422 -----GCAATAATAGGT-----CCTTATACTCTGAAAGCT 2451  
Db : : : : :  
702 GlySerGlySerAlaSerAsnSerGlyGluValThrIleThrLysHisThrLeuValVal 721  
2452 ACATGGGCTAAACTGGGTATTAATCTCGGCGCTCAGCGAGTCTTTTGGTTCCAAAT 2511  
Db : : : : :  
722 AsnTrpAlaProValGlyTyrIleValAspProIleArgGlyAspLeuIleAlaAsn 741  
2512 AGTTTATGGGATCCATTTTAGATATATAGTATCTCGGCAATTCAGCAATTCAGCAAGTGTG 2571  
Qy

742 SerLeu-----ValHisSer----- 746  
2572 GATGGCGCTCTTATTGTCGAGGATTA-----TGGGTTTCT 2607  
Db : : : : :  
747 ---GlyA-ArgAsnMetThrMetGlyLeuArgSerLeuLeuProAspAsnSerTrpPheAla 765  
2608 -----GGAGTTTTCGAATTTCTTCTATCATGACCGCGATGCTTTAGT---CAGGGA 2655  
Db : : : : :  
766 LeuGlnGlyAlaAlaThrThrLeuPheThrLysGlnGlnLysArgLeuSerTyrHisGly 785  
2656 TATCGGTATATTAGTGGGGTTTATCTCTTAGAGAGCAACTCTCTACTTTTGGATCATCGATG 2715  
Db : : : : :  
786 TyrSerSerAlaSerLysGlyTyrThrValSerSerGlnAla-----SerGlyAla 802  
2716 TTTGCTCTAGCATTTTACCGAAGTATTTGCTAGATCTTAAAGATTATGCTAGTGTCTCTTCC 2775  
Db : : : : :  
803 HisGlyHisLysPheLeuLeuSerPheSerGlnSerSerAspLysMetLysGluLysGlu 822  
2776 AATCATCATGCTTGCATAGGATCCGTTTATCTATCTATCTACCCAAACAGCTTTTATGTGATCC 2835  
Db : : : : :  
823 ThrAsnAsnArgLeuSerSerArgTyrTyrLeuSer-----AlaLeuCysPheGlu 839  
2836 TATTGTTTCGGAGAT-----CGGTTTATC---CGTGTAGCTACGGGTTTGGGAATCAG 2886  
Db : : : : :  
840 HisProMetPheAspArgIleAlaLeuIleGlyAlaAlaAlaCysAsnTyrGlyThrHis 859  
2887 CATATGAAAACCTCATATACATTTGCAGAGGAGCGATGTTGCTTTGGGATATAACTGT 2946  
Db : : : : :  
860 AsnMetArgSerPheTyrGlyThrLysLysSerSerLysGlyLysPheHisSerThrThr 879  
2947 CTGGCTGGAGAGATTCGAGCGGATTTACCGATTTGTGATTACTCTCATCT---AAGCTCTAT 3003  
Db : : : : :  
880 Leu-----GlyAlaSerLeuArgCysGluLeuArgAspSerMetProLeuArg 895  
3004 TTGAATAGTTCGCTCTTTCGCAAGCTGAGTCTTTTCTTATGCGCATCATGATCTTTT 3063  
Db : : : : :  
896 SerIleMetLeuThrProPheAlaGlnAlaLeuPheSerArgThrGluProAlaSerIle 915  
3064 ACAGGAGAGCGCATCAAGCTCGGCAATTC-----AAGAGCGGACATCTCTAAATCTA 3117  
Db : : : : :  
916 ArgGluSerGlyAspLeuAlaArgLeuPheThrLeuGluGlnAlaHisThrAlaValVal 935  
3118 TCAGTTTCTGTTGGAGTGAAGTTTGTATCGATGTTCTAGTACACATCTCTAAT---AAATAT 3174  
Db : : : : :  
936 Ser---ProIleGlyLysGlyAlaTyrSerSerAspThrTrpProThrLeuSerTrp 954  
3175 AGCTTTATGCGGCTTTATATCTGTGATGCTTATCGCACCATCTCTGTGACTGAGACACAG 3234  
Db : : : : :  
955 GluMetGluLeuAlaTyrGlnProThrLeuTyrTrpLysArgProLeuLeuAsnThrLeu 974  
3235 CTCTATCTCCATCAAGAGCATGACACAGATGCTCTTTCATTTTACGACATGGATTT 3294  
Db : : : : :  
975 LeuIleGlnAsnAsnGlySerTrpValThrThrAsnThrProLeuAlaLysHisSerPhe 994  
3295 GTGTTTAGAGGATCTATGATGCTCTCTAACAGTAATATAGATATATATGCGCATCGGA 3354  
Db : : : : :  
995 TyrGlyArgGlySer---HisSerLeuLysPheSerHisLeuLysLeuPheAlaAsnTyr 1013  
3355 AGATATGATGATCGAGATGCTTCTCGAGGCTTCTGTTTGTAGTCAGGAAGTAGATCGCG 3414  
Db : : : : :  
1014 GlnAlaGluValAlaThrSerThrValSerHisTyrIleAsnAlaGlyGlyAlaLeuVal 1033  
3415 TTC 3417  
Db : : : : :  
1034 Phe 1034

## RESULT 16

US-11-103-957-45  
; Sequence 45, Application US/11103957  
; Publication No. US20050281847A1  
; GENERAL INFORMATION:  
; APPLICANT: Berthet, Francois-Xavier Jacques

```

; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1751
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-45

Alignment Scores:
  Pred. No.:      1,19e-32      Length:      1751
  Score:          475.50      Matches:      269
  Percent Similarity: 36.6%      Conservative: 148
  Best Local Similarity: 23.6%      Mismatches: 454
  Query Match:      6.0%      Indels:      269
  DB:                11      Gaps:        51

US-10-701-844-1 (1-4435) x US-11-103-957-45 (1-1751)
QY 484 GGAATTACGATGGGAGACGTTACTGTATCATTTCCCTATCTGTTATAGGATCCG 543
DB 119 GlyIleTyrAlaLysLysAlaLysMetSerArgIleAspGlnLeuAsnIleSerGluAsn 738
QY 544 AGTGGAGCTACTGTTTTCTGCAGAGAGTTAACTAAATAAAATCTTGACATTTCTATT 603
DB 739 SerAlaThrGluIle-----GlyGlyIleCysCysLysGluSerLeuGluLeuAsp 756
QY 604 GCAGCTTTGCCCTTAAGTTGTTTTGGAACTATTAGGAGGTTTACTGTTTATAGGAGA 663
DB 757 AlaLeuValSerLeuSerValThrGluAsnLeuValGlyLys-----GluGlyGly 773
QY 664 GCACACTCGTTGACTTTCGAGAACATACGACTTCTACAAATGGGGAGCTCTAAGTAAT 723
DB 774 GlyLeuHisAlaLysThrValAsnIleSerAsnLeuLysSerGlyPheSerPheSerAsn 793
QY 724 AGCGCTGCTGAGGACTGTTTACTATTGAGGGTTTTAAAGATTATCTTTTCCAAATTGC 783
DB 794 AsnLysAlaAsnSerSerSerThr-----GlyValAlaThrThr 806
QY 784 AATTCATTACTTGGCGTACTGCTGCTGCAACGACTAATAAGGTTAGCCAGACTCCGACG 843
DB 807 AlaSerAlaProAlaAlaAlaAlaAlaSerLeuGlnAlaAlaAlaAlaValProSer 826
QY 844 ACAACATCTACACCTCTAAGTACTATTATTCTTAAACAGACTCTTTTGTACTCAAT 903
DB 827 SerProAlaThrProThr----- 832
QY 904 ANTGAGAGTTCTCATTTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCT 963
DB 833 -----TyrSerGlyValVal-----GlyGlyAlaIleTyrGly 843
QY 964 AAGAGCTTAACGGTTCAAGGAATTAAGCAAGCTTTGTCTTCCAGAAATACTGCT--- 1020
DB 844 GluLysValThrPheSerGlnCysSerGlyThrCysGlnPheSerGlyAsnGlnAlaIle 863
QY 1021 -----CAAGCTGATGGGGAGCTTGTCAAGTAGTGCACC 1053
DB 864 AspAsnAsnProSerGlnSerSerLeuAsnValGlnGlyAlaIleTyrAlaLysThr 883
QY 1054 AGTTTCTGCTATGGCTAACGAGGCTCTTATGCTTTGTA----- 1095
DB 1199 LeuHisAsnGlyThrLeuValLeuLysGlyLysThrGluLeuHisValSerPheGlu 1239

```





[illegible]





1500 ThrSerGluGluPheThrTyThrSerArgGlyAlaSerValAlaLeuAspAlaLysPro 1519  
2704 GGATCATCATGTTT---GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTAT 2760  
1520 AlaHisAspValIleValGlyAlaAlaPheSerLysMetIleGlyLysThrLysSer--- 1538  
2761 GTAGGTGTGCTTCCATCATCATGCTTGCATAGATCGTTTATCTATCTATCCCAACAA 2820  
1539 ---LeuLysArgGluAsnAsnTyThrHisLysGlySerGluTy---SerTyGlnAla 1556  
2821 GCTTATGTGGATCCCTATTGTTCCGAGATGCGTTTATCCGT----- 2862  
1557 SerValTyGlyGlyLysProPheHisPheValIleAsnLysThrGluLysSerLeu 1576  
2863 -----GCTAGCTACGGTGGTGGGAATCAGCATATGAACCTCA 2901  
1577 ProLeuLeuGlnGlyValIleSerTyGlyTyIleLysHisAspThrValThrHis 1596  
2902 TAT---ACATTTGCAGAGAGCGATGTTGTTGGATAATAACTGTCTGGCTGGAGAG 2958  
1597 TyProThrIleArgGluArgAsnLysGlyGluTrpGluAsp-----LeuGlyTrp 1613  
2959 ATTGAGCG-----GGATTACCGATTGTGATTCTCCATCAAGCTCTATTGAAATGAG 3012  
1614 LeuThrAlaLeuArgValSerValLeuArgThrProAlaGlnGlyAspThrLysArg 1633  
3013 TTGCGTCTTTCGTCGCAAGCTGATGTTTCTTATGCGCATCATCAATCTTTACAGAGGA 3072  
1634 IleThrValTyGlyGluLeuGluTySerSerIleArgGlnLysGlnPheThrGluThr 1653  
3073 GCGCATCAAGCTCGGCATTCAAGAGCGGCACATCTCTAAATCTATCAGTCTCTGTTGGA 3132  
1654 GluTyAspProArgTyPheAspAsnCysThrTyArgAsnLeuAlaIlePrometGly 1673  
3133 GTCAAGTTTGAT---CGATGTTCTAGTACACATCTCTAAATATATAGC---TTTATGGCG 3186  
1674 LeuAlaPheGluGlyGluLeuSerGlyAsnAspIleLeuMetTyArgAsnArgPheSerVal 1693  
3187 GCTTATATGTGATGCTTATCCGACCATCTCTGCTAGTACGACAGCTCTATCCAT 3246  
1694 AlaTyMetLeuSerIleTyArgAsnSerProThrCysLysTyGlnValLeuSerSer 1713  
3247 CAAGACATGGACACAGATGCTTTCATTTAGCAGACATGAGTGGTGGTTAGAGGA 3306  
1714 GlyGluGlyGlyGluIleLeuCys-----GlyValProThrArgAsn 1727  
3307 TCTATGATGCTTCTCTAACAGTAATAA-----GAAGTATATGGC 3348  
1728 SerAlaArgGlyGluTySerThrGlnLeuTyLeuGlyProLeuTrpThrLeuTyGly 1747  
3349 CATGGAAGATATCAGTATCGAGATGCTTCTCGAGGCTATGGTTGAGTCAGGAGTAGA 3408  
1748 SerTyThrIleGluAlaAspAlaHisThrLeuAlaHisMetMetAenCysGlyAlaArg 1767  
3409 GTCGGGTTTC 3417  
1768 MetThrPhe 1770

RESULT 19  
US-11-018-868-17  
; Sequence 17, Application US/11018868  
; Publication No. US20060034871A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; APPLICANT: Bonci, Alessandro  
; APPLICANT: Finco, Oretta  
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis  
; FILE REFERENCE: 002441.00099 (pp23152.001)  
; CURRENT APPLICATION NUMBER: US/11/018.868  
; CURRENT FILING DATE: 2004-12-22  
; NUMBER OF SEQ ID NOS: 186

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17  
; LENGTH: 1770  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-11-018-868-17  
Alignment Scores:  
Pred. No.: 4,19e-27 Length: 1770  
Score: 412.00 MatChes: 258  
Percent Similarity: 35.5% Conservative: 148  
Best Local Similarity: 22.6% Mismatches: 421  
Query Match: 5.2% Indels: 316  
DB: 11 Gaps: 51

US-10-701-844-1 (1-4435) x US-11-018-868-17 (1-1770)  
QY 673 TTGACTTCGAGAACATACGGACTCTTACA-----AATGGG 708  
Db 716 IleThrAspGluSerValSerSerGluSerGlySerThrProGlnAspGly 735  
QY 709 GCAGCTTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGTTTAAAGAAATTA 768  
Db 736 GlyAlaAlaSerSerGlyAlaProSerGlyAspGlnSerIle----- 749  
QY 769 TCCTTTTCCAAATGCAATTCTACTTGCCTACTGCTCTGCTGCAACGACT----- 819  
Db 750 -----SerAlaAsnAlaCysLeuAlaLysSerTyAlaAlaSerThrAspSerSer 766  
QY 820 -----AATAAGGGTAGCCAGCTCCGACGACCAACATCTACACCGTCTAATGGT 867  
Db 767 ProValSerAsnSerSerGlySerGluProValThrSerSerSerAspSerAspVal 786  
QY 868 ACTATTATTCTAAACAGATCTTTGTTACTCAATAATGAGAAGTTCTCATTTCTATAGT 927  
Db 787 ThrAlaSerSerAsp-----AsnProAspSerSerSer----- 797  
QY 928 AATTGATGCTCTGGAGAT----- 945  
Db 798 -----SerGlyAspSerAlaGlyAspSerGluGluProThrGluProGluAlaGly 814  
QY 946 -----GGGGAGCTATAGATCTAGATCTAGAGCTTAACG 975  
Db 815 SerThrThrGluThrLeuThrLeuIleGlyGlyGlyAlaIleTyArgGluThrValLys 834  
QY 976 GTTCAAGGAATTAGCAAGCTTTGTCTCTTCCAGAAAATACTGCT-----CAAGCT 1026  
Db 835 IleGluAsnPheSerGlyGlnGlyPheSerGlyAsnLysAlaIleAspAsnThrThr 854  
QY 1027 GATGGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCTAACGAGGCTCCTATT 1086  
Db 855 GluGlySerSerLys----- 860  
QY 1087 GCCTTTGTAGCGAATGTTCCAGGAGTAAAGGGGAGGAGTCTGCTGTTAGGATGGG 1146  
Db 861 -----SerAspValLeuGlyGlyAlaValTyAlaLysThrLeuPhe 874  
QY 1147 CAGCAGGGAGTGCATCTACTTCTCAACAGAGATCCAGTAGTAGTAAAGTTTCCAGAAAT 1206  
Db 875 AsnLeuAspSerGlySerSerArgThr-----ValThrPheSerGlyAsn 890  
QY 1207 ACT-----GCGGTAGAGTTTGTATGGACGCTAGCCGAGTAGGAGGAGGATTTACTCC 1260  
Db 891 ThrValSerSerGlnSerThrThrGlyGlnValAla-----GlyGlyAlaIleTySer 908  
QY 1261 -----TACGGGAACGTTGCTTCTCTGAATAATGAAACCTTTGTTTCTC 1305  
Db 909 ProThrValThrIleAlaThrProValValPheSerLysAsnSerAlaThr----- 925  
QY 1306 AACAAATGTTCTCTCTGTTTACATGCTGCTAAGCAACCAACAGTGCACAGGCTTCT 1365  
Db 926 AsnAsnAlaAsnAsnThrThrAspThrGlnArgLysAspThrPheGlyGlyAlaIleGly 945



```

Db      1634 IleThrValTyrGlyGluLeuGluTyrSerSerIleArgGlnLysGlnPheThrGluThr 1653
QY      3073 GCGATCAAGCTCGGCATTCAAGACGACATCTCTAAATCTATCAGTCTCTGTGGA 3132
Db      1654 GluTyrAspProArgTyrPheAspAsnCysThrTyrArgAsnLeuAlaIlePrometGly 1673
QY      3133 GTGAAGTTTGAT---CGATGTTCTAGTACACATCTCTAAATAATATAGC---TTTATGGCG 3186
Db      1674 LeuAlaPheGluGlyGluLeuSerGlyAsnAspIleLeuMetTyrAsnArgPheSerVal 1693
QY      3187 GCTTATATCTGTGATCTTATCGCACCATCTCTGTGACTGAGACACAGCTCTCTATCCCAT 3246
Db      1694 AlaTyrMetLeuSerIleTyrArgAsnSerProThrCysLysTyrGlnValLeuSerSer 1713
QY      3247 CAAGACATGAGCAACACATGCTTTCATTATGACGACATGAGTGTGTTAGAGCA 3306
Db      1714 GlyGluGlyGlyGluIleCys-----GlyValProThrArgAsn 1727
QY      3307 TCTATGTATGCTTCTCTAACAGTAATATA-----GAAGTATATGGC 3348
Db      1728 SerAlaArgGlyGluTyrSerThrGlnLeuTyrLeuGlyProLeuTrpThrLeuTyrGly 1747
QY      3349 CATGGAAGATATGAGTATGAGTATGAGTCTCTCGAGGCTATGTTTGAGTCGAGCAAGTACA 3408
Db      1748 SerTyrThrIleGluAlaAspAlaHisThrLeuAlaHisMetAsnMetCysGlyAlaArg 1767
QY      3409 GTCCGGTTC 3417
Db      1768 MetThrPhe 1770

```

```

RESULT 20
US-11-052-554A-1
; Sequence 1, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1268
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-1

```

```

Alignment Scores:
Pred. No.: 4,43e-15 Length: 1268
Score: 273.50 Matches: 304
Percent Similarity: 33.2% Conservative: 187
Best Local Similarity: 20.6% Mismatches: 495
Query Match: 3.5% Indels: 493
DB: 11 Gaps: 79

```

```

US-10-701-844-1 (1-4435) x US-11-052-554A-1 (1-1268)
QY      415 ATGATTTCTAGCTATTCTTGTCTCTTTAAATGGGGGGGATGATGAGCGAGAAATCATG 474
Db      21 MetMetLeuSerPheProValAlaSer-----GlnAlaAlaGlyLeuVal 35
QY      475 GTTCCTCAAGA---ATTACGATGGGAGACGTTAACTGATCATTTCCCTATCTGTT 531
Db      36 IleLysAsnGlyThrValTyrAsnAlaAsn-----GlyValProValValAsp 51
QY      532 ATAGGAGATCGAGTGGGACTGTTTCTGCGAGGAGTTAAACATTAATAAAATCTT 591

```

```

Db      52 IleAsnLysProAsnGlySerGlyLeu---SerHisAsnIleTrpAspAsnLeuAsnVal 70
QY      592 GAC-----AAATCTATTGACGCTTTTGCCTTTTAAAGTTGTTTGGG 630
Db      71 AspLysAsnGlyValValPheAsnAsnSerAlaAsnGluSerSerThrSerLeuAlaGly 90
QY      631 AACTATTATTAGGAGTTTACTGTTTATGGGAGAGACACTCGTGTACCTTCGAGACATA 690
Db      91 AsnIleGlnGlyAsn---SerAsnLeuThrSerGlySerAlaLysValIleLeuAsnGlu 109
QY      691 CGGACTTCTACAAATGGGCGACCTCTAAGTAATAGCGTGTGATGGACTGTTTACTATT 750
Db      110 ValThrSerLysAsnProSer-----ThrIle 118
QY      751 GAGGCTTTTAAAGAAATTATCTTTTCCAATTTCATTTACTTTCCTGCTGCTGCT 810
Db      119 AsnGlyMetMetGluValAlaAlaGlyAspLysAlaAspLeuIleAlaAsnProAsnGly 138
QY      811 GCAACGACTTAATAAGGTTAGCCAGACTCGAGC-----ACAATCTACA 855
Db      139 IleThrValAsnGlyGlyGlySerIleAsnThrGlyLysLeuThrLeuThrThrGlyThr 158
QY      856 CCG-----TCTAATGGTACTATT----- 873
Db      159 ProAspIleGlnAspAspLysLeuAlaGlyTyrSerValAsnGlyGlyThrIleThrLeu 178
QY      874 -----TATTCTAAACACAGATCTTTTG-----TTACTCAATAAT 906
Db      179 GlyLysLeuAspAsnAlaSerProThrGluIleLeuSerArgAsnValValValAsnGly 198
QY      907 GAGAAGTTCTCATTTCTATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAG 966
Db      199 LysValSerAlaAspGluLeuAsnValValAlaGlyAsn---AsnTyrValAsnAlaAla 217
QY      967 AGCTTAACGTTCAAGCAATTTAGCAAGCTTTGCTTCCCAAGAAATATGCTCAAGCT 1026
Db      218 GlyGlnValThrGlySerValSerAla-----ThrGlySerArg 230
QY      1027 GATGGGGAGCTGTGTCAGTAGTACCAGCTTTCTCTGCTATGCTACGACGAGCTCTATT 1086
Db      231 AsnGlyTyrSerValAspValAlaLysLeuGlyGlyMetTyrAlaAsnLysIleSerLeu 250
QY      1087 GCCTTTGTAGCGCAATTTGCAGGAGTAAGAGGGAGGAGGATGCTGCTGTTCAGGATGGG 1146
Db      251 ValSerThrGluLysGlyValGlyValArgAsnLeuGlyValIleAla-----GlyGly 268
QY      1147 CAGCAGGAGTGTC----- 1161
Db      269 ValAsnGlyValSerIleAspSerLysGlyAsnLeuAsnSerAsnAlaGlnIleGln 288
QY      1162 -----TCATCTACTTCAACAGAA 1179
Db      289 SerAlaSerThrIleAsnLeuThrThrAsnGlyThrLeuAspAsnThrThrGlyThrVal 308
QY      1180 GATCCAGTAGTAGTTTTCCTCC-----AGAAATACTCGGTAGAGTTT---GAT 1224
Db      309 ThrSerValGlyThrIleSerLeuAsnThrAsnLysAsnThrIleValAsnThrArgAla 328
QY      1225 GGGACGTCACCGAGTAGGA----- 1245
Db      329 GlyAsnIleSerThrMetGlyAspIleTyrValAsnSerGlyThrIleAspAsnThrAsn 348
QY      1246 GGAGGGTTTACTCTTACCGGAACGTTGCTTCTGTAATAATGMAAAACCTTGTCTTC 1305
Db      349 GlyLysLeuAlaAlaAlaGlyMetLeuAlaValAspThrAsnAsnAlaThrLeuIleAsn 368
QY      1306 AACAAATGTTCTTCTCTCTGTT-----TACATTGCTGCTAGCAACCAACA 1350
Db      369 SerGlyLysGlySerSerValGlyIleGluAlaGlyLeuValAlaLeuLysThrGlyThr 388
QY      1351 AGTGGACAGCTTCTTAATACGAGTAATAATTACGAGATCGAGGAGCTATCTTCTGTAG 1410
Db      389 -----LeuAsnAsnSerAsnGlyGlnIleArgGlyGlyTyrValGlyLeuGlu 404

```





```

; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1245
; LENGTH: 3132
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-11-087-099-1245

Alignment Scores:
Pred. No.: 1,22e-14 Length: 3132
Score: 270.50 Matches: 300
Percent Similarity: 34.1% Conservative: 192
Best Local Similarity: 20.8% Mismatches: 560
Query Match: 3.4% Indels: 390
DB: 11 Gaps: 67

US-10-701-844-1 (1-4435) x US-11-087-099-1245 (1-3132)
QY 505 TTAACCTGATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTTCT 564
Db 505 TTAACCTGATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTTCT 564
QY 35 IleMetLeuSerPheProGlyGlnAlaLeuAlaGlyProSerGlyGlnValIle 53
Db 35 IleMetLeuSerPheProGlyGlnAlaLeuAlaGlyProSerGlyGlnValIle 53
QY 565 GCAGGAGAGTTAAACATTA---AAAATCTTGCAATTTCTATTGCGAGCTTTGCTTTAAGT 621
Db 565 GCAGGAGAGTTAAACATTA---AAAATCTTGCAATTTCTATTGCGAGCTTTGCTTTAAGT 621
QY 54 AlaGlyGlnValThrIleAspArgAsnLeuAsnAspThrIle 67
Db 54 AlaGlyGlnValThrIleAspArgAsnLeuAsnAspThrIle 67
QY 622 TGTTTTGGGAACCTTATTAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 691
Db 622 TGTTTTGGGAACCTTATTAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 691
QY 67 67
Db 67 67
QY 682 GAGACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
Db 682 GAGACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
QY 68 ---IleAsnGlnAlaThrAspLysAlaIleAsnTrpLysAsnPheAsp 83
Db 68 ---IleAsnGlnAlaThrAspLysAlaIleAsnTrpLysAsnPheAsp 83
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATATCTCTTTTCCAAATTTGCAATTTCTATTGCTCGGTA 801
Db 742 TTTTACTATTGAGGGTTTTAAAGAAATATCTCTTTTCCAAATTTGCAATTTCTATTGCTCGGTA 801
QY 84 84
Db 84 84
QY 802 CTGCGCT---GCTCCACAGCTAATAAGGTAGCCAGACTCCGAGCAACA 849
Db 802 CTGCGCT---GCTCCACAGCTAATAAGGTAGCCAGACTCCGAGCAACA 849
QY 93 IleGlnProAspValGlySerIleAlaLeuAsnArg---ValGluSer 107
Db 93 IleGlnProAspValGlySerIleAlaLeuAsnArg---ValGluSer 107
QY 850 TCTACACCGCTCTAAT---GGTACTATTATTCTTAAACACAGACTCTTTTGTACTCAAT 903
Db 850 TCTACACCGCTCTAAT---GGTACTATTATTCTTAAACACAGACTCTTTTGTACTCAAT 903
QY 108 AsnLeuProThrAsnIleLeuGlyGlnLeuGlnAlaAsnGlyArgLeuPheIleLeuAsn 127
Db 108 AsnLeuProThrAsnIleLeuGlyGlnLeuGlnAlaAsnGlyArgLeuPheIleLeuAsn 127
QY 904 AATGAGAAGTTCTCTATTATATATTTAGTCTCTGGAGATGGGGGAGCTATAGATGCT 963
Db 904 AATGAGAAGTTCTCTATTATATATTTAGTCTCTGGAGATGGGGGAGCTATAGATGCT 963
QY 128 ---AlaAsnGlyValIlePheGlyArgSerAlaGlnValAspVal 141
Db 128 ---AlaAsnGlyValIlePheGlyArgSerAlaGlnValAspVal 141
QY 964 AAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAA 1023
Db 964 AAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAA 1023
QY 142 AlaGlyLeuLeuAlaThrThrLeuSer-----IleLysAspAsnAsnPheLeu 157
Db 142 AlaGlyLeuLeuAlaThrThrLeuSer-----IleLysAspAsnAsnPheLeu 157
QY 1024 GCTGATGGGGAGCTTGTCTCAAGTAGTACCAGTTTC-----TCTGCTATGGCTACGAG 1077
Db 1024 GCTGATGGGGAGCTTGTCTCAAGTAGTACCAGTTTC-----TCTGCTATGGCTACGAG 1077
QY 158 SerGlyLysTyrAlaPheGluLysValProGlyLeuLysAspAlaAlaValGluAsnLys 177
Db 158 SerGlyLysTyrAlaPheGluLysValProGlyLeuLysAspAlaAlaValGluAsnLys 177
QY 1078 GCTCCTATT---GCCCTTTGTAGCGAATGTTGCGAGGAGTAAGA 1116
Db 1078 GCTCCTATT---GCCCTTTGTAGCGAATGTTGCGAGGAGTAAGA 1116
QY 178 GlyGluIleHisIleGlyGluGlyGlyPheAlaPheLeuVal---AlaProGlyValArg 196
Db 178 GlyGluIleHisIleGlyGluGlyGlyPheAlaPheLeuVal---AlaProGlyValArg 196
QY 1117 GGGGGAGGAGGATGCTGCTGTTTCAGGATGGGACAGGAGTGTCTATCTACTTCAACA 1176
Db 1117 GGGGGAGGAGGATGCTGCTGTTTCAGGATGGGACAGGAGTGTCTATCTACTTCAACA 1176
QY 197 AsnGluGlyLeuIleAlaLysLeuGlySerValAlaLeuAlaSerGlyAsp----- 214
Db 197 AsnGluGlyLeuIleAlaLysLeuGlySerValAlaLeuAlaSerGlyAsp----- 214
QY 1177 GAAGATCCAGTAGTAAGTTTTCAGAAATATCTGCGGTAGAGTTGATGGGAGGAGTAGCC 1236
Db 1177 GAAGATCCAGTAGTAAGTTTTCAGAAATATCTGCGGTAGAGTTGATGGGAGGAGTAGCC 1236
QY 215 -----LysMetThrIleAspPheHisGlyAsp----- 223
Db 215 -----LysMetThrIleAspPheHisGlyAsp----- 223

1008 -----GlyIleValGlyLysGlyGlyValThrLeuSerAlaGlnAsnValThrAsnAsn 1025
QY 3265 GATGCTTTTCAATTAGCAAGACATGGA-----GTTGTGGTTAGAGGATCTATG--- 3312
Db 1026 AsnSerIleIleAlaGluAsnGlyProLeuAsnLeuLeuSerArgGlyThrLeuAsp 1045
QY 3312 ----- 3312
Db 1046 AsnThrArgAlaLeuLeuSerSerGlyAlaAspAlaIleIleArgAlaAlaGlyThrPhe 1065
QY 3313 -----TATGCTTCTCTCAACAGT-----AATATAGAAGTATAT----- 3345
Db 1066 TyrAsnAsnTyrAlaThrThrThrSerAlaGlyAsnLeuAspValThrAlaAlaSerLeu 1085
QY 3346 -----GGCCATGGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTGAGT 3396
Db 1086 AsnAsnAlaSerAspGlyArgLeuGluAspAsnThrAlaThr-----GlyValIle 1102
QY 3397 GCAGGAGTAGAGTCGGTCTTCAAAAATATTGTTAGATAGTTAAGTGTAGCGATGCCT 3456
Db 1103 AlaSerAspLysAsn-----LeuAsp-----LeuSerValAsp 1113
QY 3457 TTTTCTTTTCAGATCTACATCATTTTCTTTTGTGTTTCTTCTTCTTCTTCTTCTGATGA 3516
Db 1114 AsnSerValThr-----AsnTyrGly 1120
QY 3517 TTGCGAGCTCTCTCAAGTGTTAACGCCTAATGTAAACCATCTCTTTTAAAGGAGACGAT 3576
Db 1121 TrpIleSerGly-----LysGlyAsp--- 1127
QY 3577 GTTTACTGTAATGGAGACTGCGCTTTTGTCTCAATGCTATGTCAGGAGCTGAAGAAGTTCG 3636
Db 1128 -----ValHisPheAsnValLeuLysGlyThr 1136
QY 3637 ATTATCTCAGCTAAT---GGCGACAAT---TTAACGATATTACCGACACAAACCAT 3684
Db 1137 LeuTyrAsnArgAsnAlaIleAlaIleAspAsnAlaLeuThrIleAsnAlaLeuAsnGly 1156
QY 3685 ACATTATCATTTACAGATTTCTCAAGGCCAGTTCTTCCAAAATTTATGCTTCAATTTCAGCA 3744
Db 1157 ValGluAsnPheLysAsp-----IleValAla 1165
QY 3745 GAGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAATGTTTCTTGC 3804
Db 1166 GlyThrAlaLeuThrIleAspThr-----GlnLysTyrValThrAsn 1179
QY 3805 GGAGAAAAGGAATGATCTCCGGGAAAACCGTGAGTATTTCCGAGCGACGCGAAGTGATT 3864
Db 1180 AsnSerAsnSerAsnMetLeuGlyGlnThrIleAlaIleAsnAlaValAsnAspIleAsn 1199
QY 3865 TTCTGGGATAACTCCGTTGGGTATTTCTCTTTTATCTATCTGTCGCAACCTCATCAACT 3924
Db 1200 AsnArgGlyAsnIleValGlyAspTyrSerLeuGlyValLysThr----- 1215
QY 3925 CCGCTGCTCCACAGTAGTATGTCGGAAGAAGGGTCTATTTTCTGTAGAGACTAGT 3984
Db 1216 -----GlyAsnIleTyrAsnTyrLeuAsnMet 1224
QY 3985 TTGCGAG-----ATCTCAGCGCTCAAAAAGGGTCTATGTTCCGATAATAT----- 4029
Db 1225 LeuSerTyrGlyValAlaGlyValSerAlaAsnLysValThrAsnSerGlyLysAspAla 1244
QY 4030 -----GCCGGGAATTTCCGACACAGATTTTTCGAGGTAAGAATAATAATATGCTGGT 4080
Db 1245 ValLeuGlyGlyPheTyrGlyLeuAlaLeuGluAlaAsnGluThrAspAsnThrGly 1263

RESULT 21
US-11-087-099-1245
; Sequence 1245, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
```



Db 871 ThrArgGlyGlySerIleGlyAspAlaThrHisAlaLeuYThrGluLeuGlyIleLeu 890  
Qy 3109 CTAATCTATCGATTCTGTTGGAGTGAAGTTTGTGATCGATGTTCTAGTACACATCTCTAAT 3168  
Db 891 ThrAlaSerSerSerAsnGlySerValTyrIleAspGluAlaSerGlyIleLeuIleAsn 910  
Qy 3169 AAATATAGCTTT-----ATGGCGGCTTATATCTGTGATGCTTATCGCACCATC 3216  
Db 911 SerIleAsnAlaAsnGlnIleGlyValThrAlaIleValAsnSerSerTyrAlaValVal 930  
Qy 3217 TCTGTTACTGAGACACGCTCTCTATCCATCAAGAGACATGCAACACAGATCCCTTTCAT 3276  
Db 931 -----ValThrThrLeuAspGlyGlnThrGluThrGlyThrAsnAsnValThrVal 947  
Qy 3277 TTAGCAAGACATGGAGTTGTTGAGGATCTATGTATGCTTCTTCAACAAGTAAATA 3336  
Db 948 ThrAlaAla---GlyAspIleIleIleGlyThrAlaSerAlaThrSerThrValThrLeu 966  
Qy 3337 GAAGTATATGGCCATCGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGA 3396  
Db 967 ThrSerThrAlaGlyAsnIleTyrAspAsnAsnAspAlaThrAsn----- 981  
Qy 3397 GCAGGAAGTAGAGTCGGTCTTCTAAATAATGCTTAGATAGTAAAGTGTAGCGATGCT 3456  
Db 981 ----- 981  
Qy 3457 TTTTCTTTCAGATCTACATCATTTTGTGTTTTCAGTGTGTTGTTCTTTCGTATGGA 3516  
Db 981 ----- 981  
Qy 3517 TTTCGGAGCTCTCTCAAGTGTAAAGCCCTAATGTAACCACTCTCTTTTAAGGAGGAGAT 3576  
Db 982 -----AsnIleLeuAlaIleValVal----- 989  
Qy 3577 GTTACTTCAAGTGGAGTGGCTTTTGTCAATGCTCTATGTCAGGAGCTGAAGAAGTTCG 3636  
Db 990 -----LeuAsn-----AlaGlyLeuAsnMetGly--- 997  
Qy 3637 ATTATCTAGTAATGCGACAAATTTAAGATTACCGGACAAACCATCATTCATTCATTT 3696  
Db 998 -----LeuSerGlyAspValLeuLeuLeuThrGluSerAlaThrLeuThrAla 1014  
Qy 3697 ACAGATTCTCAAGGCGCAGTCTTCAAAATATTGCTTCATTTTCAGCAGGAGACATTT 3756  
Db 1015 ThrAsn-----GlyGlyValTyrPheAlaMetGlyValProSerThrValAlaSerVal 1032  
Qy 3757 ACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAATGTTTCTTTCGCGAGAAAAGGGA 3816  
Db 1033 ThrAlaGlyGlyHisAspValValIleAsnAlaSerAspGlyValLeuSerLeuGly 1052  
Qy 3817 ATGATCTCC-----GGGAAACCGTGAGTATTTCCGGAGCAGCGGAAGTGTATTCGG 3870  
Db 1053 MetValSerAlaValGlySerValThrIleAspGlyGlyThrGlySerIleLeuAsp 1072  
Qy 3871 GATACTCCGTTGGGTATTCCTTCTTATCTGTCGCAACCTCATCATCACTCCGCT 3930  
Db 1073 AsnAsnGlyAlaThrAlaAlaIleIleAlaAsnSerAlaSerIleIleYSerAspGly 1092  
Qy 3931 GCTCCAACTAGTATGATGTCGGAAGGCTCTATTTTCTGATAGAGACTAGTTGGAG 3990  
Db 1093 AlaGlyThrAlaSerAspAlaLeuGluThrThrValAspThrLeu-----AlaValGlu 1110  
Qy 3991 ATCTCAGGCTCAAAAGGGGTGATGTTTCGATAAT-----AAT 4029  
Db 1111 IleThrGlySerGlyIleYSerPheTyrIleAspGluSerAspAlaLeuThrSerIleAsn 1130  
Qy 4030 GCGGGGAATTCGGAACAGTTTTTCGAGGTAAGATAATATATGCTGCTGCTGGAGGC 4089  
Db 1131 Ala-----LysValAsnAsnGlySerThrAsnLeuAsnPheThrGlyGlySer 1146  
Qy 4090 AGTGGGTTCCGCTACCATCAAGTACGACTTTTACAGTTTAAACACTGTAAGCGGAAGT 4149  
Db 1147 PheAlaPheAsnAlaThrThr-----GlyAlaPheSer-SerThrGlyValGlySerVa 1164

Qy 4150 TTCTTTACAGATAACGTAGCTCTTTCGGAGGCGGAGTGGTTTATAAAGCATTTGTGCT 4209  
Db 1164 lThrPheGluAsnThr-----GlyGlyGlyValAlaIleGlyThrValThrAl 1180  
Qy 4210 TTCAAAGACAAATGAAGGAGCATATTTCTCCGAGGGAACACAGCATACGATGATTAA 4269  
Db 1180 a-----ThrGlyGlySerAlaThrIleThrAlaSerThrAlaIleThrAsp----- 1195  
Qy 4270 GATTTCTCTCTACTAATCAGGAT-----CAGAAATACGAGACAGGAGG 4314  
Db 1196 -AlaThrSerAlaIleThrAlaAspThrValValLeuThrAlaGlyThrSerIleGlyAl 1215  
Qy 4315 CGGTGGAGGAGTATTTCTCTCCAGATCATTTCTGAAAGTTTGAAGGCAATAAAGGTTTC 4374  
Db 1215 aserGlySerThrIleYSerThrAlaThrThrLeuThrTyrLeuAlaSerAlaGlySe 1235  
Qy 4375 TATT 4378  
Db 1235 rfile 1236  
RESULT 22  
US-11-052-554A-91  
; Sequence 91, Application US/11052554A  
; Publication No. US20050288866A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; PRIOR FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 91  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori J99  
US-11-052-554A-91  
Alignment Scores:  
Pred. No.: 1,61e-13 Length: 2902  
Score: 257.50 Matches: 302  
Percent Similarity: 31.4% Conservative: 171  
Best Local Similarity: 20.1% Mismatches: 472  
Query Match: 3.3% Indels: 560  
DB: 11 Gaps: 79  
US-10-701-844-1 (1-4435) x US-11-052-554A-91 (1-2902)  
Qy 403 TTCTTTCTTTCAATGATCTTAGCTTATCTTCTGCTCTTTTAAATGGGGGATATGCA 462  
Db 119 TyrPheLeuSerGlyLeuTyrAsnTyrThrGlyGlyLeuTyrAsnGlyGlyAsnLeuAsp 138  
Qy 463 GCAGAAATC-----ATGGTTCCTCAAGGAATTTACGATGGGAGACGTTAACT 510  
Db 139 IleGluLeuGlySerAsnAlaThrPheAsnLeuGlyAlaSerSerGlyAsnSerPheThr 158  
Qy 511 GTATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACT-----GTTTTTCT 564  
Db 159 SerTrpTyr-----ProAsnGlyHisThrAspValThrPheSer 171  
Qy 565 GCAGAGAGTTAACTATTAATAAATCTTGACAAATCTTATTCAGCTTTGCTTAAAGTTGT 624  
Db 172 AlaGlyThrIle-----AsnValAsnAsnSerValGluVal----- 183  
Qy 625 TTTCGGGAACCTTATAGGAGTCTTACTGTTTATAGGAGAGA---CAC----- 669  
Db 184 ---GlyAsnArgValGlySer-----GlyAlaGlyThrHisThrGlyThrAla 198





```

; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 90
; LENGTH: 3194
; TYPE: PRT
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-90

```

```

Alignment Scores:
Pred. No.: 1-7e-12 Length: 3194
Score: 246.00 Matches: 327
Percent Similarity: 30.8% Conservative: 196
Best Local Similarity: 19.3% Mismatches: 594
Query Match: 3.1% Indels: 581
DB: 11 Gaps: 81

```

US-10-701-844-1 (1-4435) x US-11-052-554A-90 (1-3194)

```

QY 421 CTAGCTTATTCGTCTTTAAATGGGGGGATATGCAGCAGAAATCATGTTCT 480
DB 231 IleThrTyrAsnGlySerValAsnGlyGlyAsnPheGlyPheAspAsnValAspSer 250
QY 481 CAGGAATTCAGTGGGAGACGTTAACTGTATCATTTCCCTATCTGTTATAGAGAT 540
DB 251 AsnGlyAla-----ThrThrIle----- 256
QY 541 CGAGTGGGACTACTGTTTTCTCGAGGAGATTAAACATTAA- 585
DB 257 ---SerGlyValThrPheAsnAsnAsnGlyAlaLeuThrTyrLysGlyGlyAsnGlyIle 275
QY 586 -----AATCTTGACAAT 597
DB 276 GlyGlySerIleThrPheThrAsnSerAsnIleAsnHisTyrLysLeuAsnLeuAsnAla 295
QY 598 TCATTGCGAGCTTCCTTTAAGTTGTTTGGG-----AACTTA 636
DB 296 AsnSerValThrPheAsnAsnSerThrLeuGlySerMetProAsnGlyAsnAlaAsnThr 315
QY 637 TTAGGGAGTTTACTGTTTGGGAGAGACACTCGTTGACTTTCGAGAACATACGGACT 696
DB 316 IleGlyAsnAlaTyrIleLeu---AsnAlaAsnAsnIleThrPheAsnAsnLeu----- 332
QY 697 TCTACAAATGGGCGAGCTCTAAGTAATAGCGTGTGATGGACTGTTTACTATTAGGGT 756
DB 333 ThrPheAsnGlyGlyTyrPheValPheAsnAsnSerAspAlaHisValAsnPheGlnGly 352
QY 757 TTTAAGAATATACC-----TTTTCCAATTGCAATTCATTCTTCGCGTA 801
DB 353 ThrThrThrIleAsnAsnProThrSerProPheValAsnMetThrGlyLysValThrIle 372
QY 802 CTGCTGCTGCAACGACTAATAAGGTAGCGAGCTCCGACG- 843
DB 373 AsnProAsnAlaIlePheAsnIleGlnAsnTyrThrProThrIleGlyAsnAlaTyrThr 392
QY 843 ----- 843
DB 393 LeuPheSerMetLysAsnGlyAsnIleAlaTyrAspAspValAsnAsnLeuTrpAsnIle 412
QY 844 -----ACAACATCTACACCGTCTAAT----- 864
DB 413 IleArgLeuLysAsnThrGlnAlaThrLysAspAsnSerLysAsnAlaThrSerAsnAsn 432
QY 865 -----GGTACTATTATTCT----- 879
DB 433 AsnThrHisThrTyrTyrValThrTyrAsnLeuGlyGlyThrLeuTyrHisPheArgGln 452

```

```

QY 880 -----AAACACAGATCTTTTGTCTACTCAATAATGAGAGTTCTCA-----TTC 921
DB 453 IlePheSerProAspSerIleValLeuGlnSerValTyrGlyAlaAsnAsnLeuTyr 472
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACG----- 975
DB 473 TyrThrAsnSerValAsnIleHisAspAsnValPheAsnLeuLysAsnIleAsnAspAsp 492
QY 976 -----GTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAAAGAAAAT 1014
DB 493 ArgAlaAspThrIlePheTyrLeuAsnGlyLeuAsnThrTrpAsnTyrThrGlnAlaArg 512
QY 1015 ACTGCTCAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAAC 1074
DB 513 PheAlaGlnThrTyrGlyGlyLysAsnSerAlaLeuVal---PheAsnAlaThrThrPro 531
QY 1075 GAGGCTCTATTTGCTTTGTAGCGAATGTTGCAGAGTAAGAGGGGGAGGATTCGTCT 1134
DB 532 TrpAlaAsnGlyAlaIleProLysSerAsnSerThrValArgPheGlyGlyTyrGlnGly 551
QY 1135 GTTCAGGATGGGCGAGGAGTGTCACTCTACTTCAACAGAGATCCAGTAGTAAGT 1194
DB 552 ValAsnTrpGlyLysThrGlyTyrIleThrGlyThrPheThrAlaAspArgValTyrIle 571
QY 1195 TTTTCCAGAAATCTCGGTAGAGTTGATCGGAACGTAGCCCGAGTAGGAGGG----- 1251
DB 572 -----ThrGlyAsnMetSerGlyAsnGlyAlaGlnThrGlyGlyAla 587
QY 1252 ATTTACTCTACGGGAACGTTTCTCTGAATAATGGAACCTTCTTT-----CTC 1305
DB 588 ThrLeuAsnPheValGlyAlaThrGluIleAsnIleAlaGlyAlaThrPheLysAsnLeu 607
QY 1306 AACAAATGTTGTTCTCTTACATTTTACTGCTCTAAGCAACCAACAGAGTGGACAGCTTCT 1365
DB 608 LysThrThrSerGlnAsnSerTyrMetThrPheMetAlaLeuGlyAsnGlySerGlySer 627
QY 1366 AATACGAGTAAT-----AATTACGGAGATGGAGGAGTATCTTCT 1404
DB 628 GlyIleAlaAsnValSerGlnSerAspPheTyrAspTrpThrAspGlyGlyTyrAspPhe 647
QY 1405 TGTAAGATGTCGCGACAGCAGGATCCAAAT---AACTCGGATCAGTTTCTTCTGATCGA 1461
DB 648 ThrGlyAsnGlyValPheAspSerValAsnPheAsnIleAlaTyrTyrLysPheGlnGly 667
QY 1462 -----GAGGAGTAGTTTCTTTAGTACCAATGTAGTCTGGG-----AAA 1503
DB 668 AlaGluAsnSerTyrAsnPheLysAsnThrAsnPheLeuAlaGlyAsnPheLysPheGln 687
QY 1504 GGGGAGCTATTTATGCCAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACCAATT--- 1560
DB 688 GlyIleThrThrIleGluLys-----SerValLeuAsnAspAlaSerTyrAlaPheAsp 705
QY 1561 ---TTAAGGAATATCGCTAATGAT-----GCTGGAGCGATTATTATTAGAGAA 1605
DB 706 GlyValAsnAsnAlaPheAsnGluAspLysPheAsnGlyGlySerPheAsnPheAsnHis 725
QY 1606 TCT-----GGAGAGCTCAGTTTATCT 1626
DB 726 AlaGluGlnThrAsnAlaPheAsnAsnAsnSerPheSerGlyGlySerPheSerPheAsn 745
QY 1627 GCT-----GATTATGGAGATATTATTTCGATGGG-----AATCTTAAAGA 1668
DB 746 AlaLysGlnValAspPheAsnGlyAsnSerPheAsnGlyGlyValPheAsnPheAsn 765
QY 1669 ACAGCCAAA-----GAGAAATGCTGCCGATGTTAAT 1698
DB 766 ThrProLysAlaSerPheThrAsnAspThrPheAsnValAsnAsnGlnPheLysIleAsn 785
QY 1699 GCGCTAACTGTGTC-----TCACAAGCAATTTTCGATGGGATCGGAGGAGAAATA 1749
DB 786 GlyAlaGlnThrAspPheThrPheSerLysGlyValValPheAsnMetGlnGlyLeuLeu 805

```





Db 1351 AsnGlyGlySerIleThrPheGlyLysAsnAsnLeuLeuTyrLeuHisGlyAsnPheAsn 1370  
Qy 3601 TTGTCTCAATGCTCTAT-----CGAGGAGCTGAAGAGGT 3633  
Db 1371 AlaThrAsnIleLeuLeuThrAsnAsnPheAsnValGlyAsnProAsnAlaGlyGlyGly 1390  
Qy 3634 TCATTATCTCAGCTAATGCGGAC---AAATTAACGATTACCGGACAAACACATACATTA 3690  
Db 1391 AlaThrIleAsnPheAsnAlaAspGluThrLeuSerAlaAspGlyLeuAsnTyrThrAsn 1410  
Qy 3691 TCATTACAGATTCTCAAGGCCAGGTCTTCAA-----AATTATGCTTCATT 3738  
Db 1411 PheGlnThrValAlaMetGly-----LeuGlnThrSerAlaSerGlnHisSerTrpAla 1428  
Qy 3739 TCAGCAGGAGACACTTACTCTG-----ACAGATTCTTCG--- 3774  
Db 1429 AsnPheAsnSerLysLeuSerMetGluIleLysAsnSerAsnPheArgAspPheThrTrp 1448  
Qy 3775 ---AGTCTGATGTTCTCGAAAATGTTCTTCGGGAGAAAAGGAATGATCTCCGGAAA 3831  
Db 1449 GlyGlyPheArgPheAsnSerGlyArgIleThrPheGluAsnThrThrPheSerGlyTrp 1468  
Qy 3832 ACCGTGAGTATTCGGCAGCAGCGAA----- 3858  
Db 1469 Thr---AsnIleAsnGlyAlaThrGluSerGlySerSerTyrValAsnMetValAlaAsn 1487  
Qy 3859 ---GTGATTCTTCGGGATACTCCGTGGG----- 3885  
Db 1488 ThrAspLeuIlePheThrAspSerIleLeuGlyGlyGlyIleArgTyrAspLeuLysAla 1507  
Qy 3885 ----- 3885  
Db 1508 AsnAsnIleIlePheAsnAsnThrGlnMetValValAspValSerLysAsnValAsnGln 1527  
Qy 3886 -----TATTCTCTTATCTACTGTGCGCAACTCA 3915  
Db 1528 SerSerLeuAsnGlyAsnValThrPheAsnHisSerArgLeuSerValLysProAsnAla 1547  
Qy 3916 TCA-----TCACCTCCCGCTGCTCCACAGTTAGTGCTCGAAGGCTCT 3963  
Db 1548 AlaIleAsnIleGlyIleAspGlnThrGlnThrThrLeuGluAsnAlaSerSerLeuSer 1567  
Qy 3964 ATTTTCTCTGAGAGACTAGTTTGGAGATCTCAGCGCTCAAA-----AAAGGGCTCATG 4017  
Db 1568 PheTyrAsn---AspSerValAlaAsnPheAsnGlyThrAlaPheAsnGlyValSer 1586  
Qy 4018 TTCGATAATATATCGCGGAATTCGGACAGTTTTCGAGTAAGAATAATAATATGCT 4077  
Db 1587 TyrLeuAsnLeuAsnProAsnAlaGlnValSerPheAsnGlnAlaAsnPheAsnAsnAla 1606  
Qy 4078 GGTGCTGAGCAGCTGGT-----TCCGCTACACATCA---AGTACGACTTT 4122  
Db 1607 Asn-ValThrPheTyrGlyIleProLeuPheGlyLysThrProAsnPheGlyAsnSerVal 1626  
Qy 4123 TACAGTAAAACGTAAAGGCAAGTTCTTTC-----ACAGATACTAGCTC 4173  
Db 1626 lArgLeuIleAsnPheLysGlyAspAlaLysPheAsnGlnAlaThrLeuAsnLeuArgAl 1646  
Qy 4174 TTCGGAGCGCGAGTGTTTATAAGGCATTTGTGCTTTTCAAAGCAATGAAGGA---GG 4230  
Db 1646 aLysAsnIleHisLeuAsnPheGlnGlyAlaSerThrPheGluAsnAsnSerThrMetAs 1666  
Qy 4231 CATATTCTTCGGGGAACACAGCATACGATGATTAAAGATTCTTGCTGCTACTATCA 4290  
Db 1666 nLeuAlaGluSerSerGlnAlaSerPheAsnAlaLeuSerValGluGlyGluThrAsnPh 1686  
Qy 4291 GGATCAGAAATACGGAG-----ACAGAGCGCGTGGAGAGTTATTTCCTCTCCAGA 4341  
Db 1686 asnLeuAsnGlySerSerLeuLeuSerPheAsnGlyAsnSerValPheAsnAlaPro-- 1705  
Qy 4342 TGATTCTTAAGTTTGAAGGCAATAAAGGTTCTATTGTTTTTGTATACAAC 4393  
Db 1706 -----ValAsnPheTyrAlaAsnAsnSerGlnIleSerPheThrHisSer 1720

## RESULT 24

US-11-052-554A-16  
; Sequence 16, Application US/11052554A  
; Publication No. US2005028866A1  
; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 16  
; LENGTH: 1250  
; TYPE: PRT  
; ORGANISM: Escherichia coli 0157:H7  
US-11-052-554A-16

Alignment Scores: 1.35e-12 Length: 1250  
Pred. No.: 245.00 Matches: 273  
Score: 245.00 Conservative: 144  
Percent Similarity: 31.1% Mismatches: 385  
Best Local Similarity: 20.3% Indels: 540  
Query Match: 3.1% Gaps: 65  
DB:

US-10-701-844-1 (1-4435) x US-11-052-554A-16 (1-1250)

Qy 553 ACTGTTTTTCTGCAGAGAGTTAAACATTAATAAATCTTGCACATTTCTATTGCAGCTTTG 612  
Db 20 SerLeuPheSerAla-----AsnGlyValAlaAlaAla 30  
Qy 613 CCTTTAAGTTGTTTGGAACTTTATTAGGAGTTTACTGTTTATAGGAGGAGGACACTCG 672  
Db 31 IleAspLeuCysGlnGly-----TyrAspIleLysAlaSerCysHisAla 45  
Qy 673 TTGACTTTTCGAGAACATACCGACTTCTCAAAATGGGCGAGCTTAAGTAATAGCGTCT 732  
Db 46 -----SerArgGlnSerLeuSerGlyIleThrGlnValTrpSerIleAla 60  
Qy 733 GATGGACTGTTTACTATTGAGGTTTTTAAGAATTATCTTTTCCAATTGCAATTCATTA 792  
Db 61 AspGlyGlnTrp-----LeuValPheSerAsp----- 69  
Qy 793 CTTGCCGTACTGCTGCTGCAACGACTAATAAGGGTAGC----- 831  
Db 70 -----MetThrAsnAsnAlaSerGlyGlyAlaValPheLeuGln 82  
Qy 832 CAGACTCCGACGACACATCTACACCGTCTAATGGTACTATTATTCTTAAACAGATCTT 891  
Db 83 GlnGlyAlaGluPheThrLeuSerProGluAsn-----GluThrGlyMet 97  
Qy 892 TTGTTACTCAATAATAGAGAGTTCTCATCTATAGTAATTATTTAGTCTCTGGAGAT----- 945  
Db 98 ThrLeu-----PheAlaAsnAsnThrValSerGlyGlyTrpAsn 110  
Qy 946 ---GGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGTC 1002  
Db 111 AsnGlyGlyAlaIlePheAlaLys----- 118  
Qy 1003 TTCCAAGAAAATATGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTAGTACCAGTTTCTCT 1062  
Db 119 -----GluAsnSerThr----- 122  
Qy 1063 GCTATGCTTAACAGGCTCCTATTTCCTTTTGTAGCAATGTTGCAGGAGTAAGAGGGGA 1122  
Db 123 -----LeuAsnLeuThrAspValIlePheSerGlyAsnValAlaGlyGlyTyrglyGly 140

```
QY 1123 GGGATTGCTGCTTTCAGGATGGCGAGGAGGTGCATCATCTACTTCAACAGAGAT 1182
Db      |||
141 AlaIle-----TyrSerSerGlyThrAsnAsp 149
QY 1183 CCAGTAGTAAGTTTTCCAGAAATATCGCGTAGAGTTTGTAGGAAACGTAGCCCGA--- 1239
Db      |||
150 ThrGlyAlaIleAspLeuArgValThrAsnAlaValPheArgAsnAsnIleAlaAsnAsp 169
QY 1240 ---GTAGGAGGAGGATTACTCCCTACGGGAACGTTGCTTCTCTGAAATATGAAACACC 1296
Db      |||
170 GlyIleGlyAlaIleThrIleAsnAsnAspIleThrLeuSerAsp-----Asp 187
QY 1297 TTGTTTCTCAACAAGTTGCTTCTCTGTTTACATGCTGCTTAAGCAACCAAGTGA 1356
Db      |||
188 ValPheAsnAsnGlnAla----- 194
QY 1357 CAGGCTTCTTAATACGAGTAATAATACGAGATGGAGGAGCTATCTTCTGTAAGAATGGT 1416
Db      |||
195 ---TyrThrSerThrSerTyrSerAspGlyAspGlyAlaIle----- 208
QY 1417 GCGAAGCAGGATCCAAATCTCGATCAGTTCCTTCTGATGGAGGAGGATGTT--- 1473
Db      |||
209 ---AspValThrAspAsnAsnSerAspSerLysHisProSerGlyTyrThrIleIleAsn 227
QY 1474 -----TTCTTTAGTAGCAATAGCTGCTGCGMAAGGGGAGCTATTTATGCCAAAAG 1527
Db      |||
228 AsnThrAlaPheThrAsnAsnThrAlaGluGlyTyrGlyAlaIleTyr----- 244
QY 1528 CTCTCGTTGCTCAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATGAT--- 1581
Db      |||
245 -----ThrAsnSerAlaThrAlaProTyrLeuIleAspIleSerValAspAspSer 261
QY 1581 ----- 1581
Db 262 TyrSerGlnAsnGlyValLeuValAspGluAsnAsnSerAlaAlaGlyTyrGlyAsp 281
QY 1582 -----GCTGGAGCATTTATTTAGGAGATCTCGAGAGCTCAGCTTTA 1623
Db 282 GlyProSerSerAlaAlaGlyPheMetTyrLeuGlyLeuSer---GluValThrPhe 300
QY 1624 TCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACACAGCAAGAGAAAT 1683
Db      |||
301 AspIleAlaAspGlyLysThrLeuValIleGlyAsn-----ThrGluAsnAspGly 317
QY 1684 GTCGCGCATTTAATGGCGTAATGCTGCTCTCAAGCCATTTTCGATGGAGTGGAGGG 1743
Db      |||
318 AlaValAsp-----SerIleAlaGlyThrGly 326
QY 1744 AATAACGACATTAAGAGCTAAAGCAGGCGATCAGATTCTC----- 1785
Db      |||
327 LeuIleThr-----LysThrGlySerGlyAspLeuValLeuAsnAlaAspAsnAsnAsp 344
QY 1786 TTTAATGATCCCATCAGATGGCAACGGAATAACACGAGCAGCTCTTCCAAACTT 1845
Db      |||
345 PheThrGlyGluMetGlnIleGluAsnGlyGluValThrLeuGlyArgSerAsnSerLeu 364
QY 1846 CTAATAAATTAAACGATGGTGAA-----GGATACACAGGG 1878
Db      |||
365 MetAsnValGlyAspThrHisCysGlnAspAspProGlnAspCysTyrGlyLeuThr--- 383
QY 1879 GATATTGTTTCTTAATGGAACGATCTTCTGATCCAAATGTTACGATAGCAGGAGGA 1938
Db      |||
384 -----IleGlySerIleAspLysTyrGlnAsn----- 392
QY 1939 AGGATTGTTCTTCGTGAAGAGGCAAAATATCAGTGAATCTCTAAGTCAGACA----- 1992
Db      |||
393 -----GlnAlaGluLeuAsnValGlySerThrGlnGlnThrPheAla 406
QY 1993 -----GGTGGAGTCTCTATATATGGAAGCTGGGAGTACATGGGAT 2031
Db      |||
407 HisSerLeuThrGlyPheGlnAsnGlyThrLeuAsnIleAspAlaGly----- 422
```

```
Db      637 ValAspAlaGly-:::--ValAspThr 643
Qy      3124 CCTGTGAGTGAAGTTTGATCGATGTTCTAGTACACATCCTAATAATATAGCTTATG 3183
Db      644 GlnTrpGlyAlaLeuMetAlaAspSerSerGlyGlnHisGlnAspGluGlySerThrLeu 663
Qy      3184 GCGGCTTATATCTGTGATGCTTATGCG--ACCATCTCTGTGACTGAGACAAACGCTCCTA 3240
Db      664 ThrLysThrGlyAlaGlyThrLeuGluLeuThrAlaSerGlyThrGlnSerAlaVal 683
Qy      3241 TCCCATCAAGACACATG--ACACAGATGCTTTTCATTTAGCAGACAT 3288
Db      684 ArgValGluGluGlyThrLeuGlnGlyAspValAlaAspIlePheProTyrAlaSerSer 703
Qy      3289 GGAGTTGTGTTAGAGGATCTATGATCTCTCTAACA-----AGTAATATA 3336
Db      704 LeuTrpValGlyAspGlyAlaThrPheValThrGlyAlaAspGlnAspIleGlnSerIle 723
Qy      3337 GAAGTATATGCGCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGTTGATG 3396
Db      724 AspAlaThrSerSerGlyThrIleAspIleSerAspGlyThr----- 737
Qy      3397 GCAGGAATAGATGCTCGGTTCTAAATAATTTGGTTAGATGTTAAGTGTAGCGATGCT 3456
Db      738 -----ValLeuArg 740
Qy      3457 TTTCTTTGAGATCATCATCTATTTTGTGTTTGTAGCTTGTG-----TGTCTCTATTCG 3510
Db      741 LeuThrGlyGlnAspThrSerValAlaLeuAsnAlaSerLeuPheAsnCysAspGlyThr 760
Qy      3511 TATGATTCGCGAGCTCTCTCAAGTGTATAGCCTTAATGTAACCACTCTCTTTTAAAGGA 3570
Db      761 LeuValAsnAlaThrAspGlyValThrLeuThrGlyGluLeuAsnThrAsnLeuGluThr 780
Qy      3571 GACGATGT--TACTGTAATGGACACTGCGCTTTTGTCTAATGTCTATGACGAGCTGAA 3627
Db      781 AspSerLeuThrTyrLeuSer----- 787
Qy      3628 GAAGGTTGCAATTATCTCAGCTAATGGCGACAAATTAACGATTAACGGGACAAACCATACA 3687
Db      788 -----AsnValThrValAsnGly----- 793
Qy      3688 TTATCATTTACAGATTCACAGGCGCAGTT--CTTCAAAATTATGCTTCAATTCACGA 3744
Db      794 ---AsnLeuThrAsnThrSerGlyAlaValSerLeuGlnAsnGlyVal-----Ala 809
Qy      3745 GGAGAGACACTTACTCTGAGA--GATTTTCG-----AGTCTGATGTTCTCGAAA 3792
Db      810 GlyAspThrLeuThrValAsnGlyAspTyrThrGlyGlyGlyThrLeuLeuLeuAspSer 829
Qy      3793 AATGTT-----TCTTGGCGAGAAAGGAATGATCTCGGGAACCGTG--- 3837
Db      830 GluLeuAsnGlyAspAspSerValSerAspGlnLeuValMetAsnGlyAsnThrAlaGly 849
Qy      3838 -----AGTATTTCCGAGCAGCGGAA----- 3858
Db      850 AsnThrThrValValAsnSerIleThrGlyIleGlyGluProThrSerThrGlyIle 869
Qy      3858 ----- 3858
Db      870 LysValValAspPheAlaAspProThrGlnPheGlnAsnAsnAlaGlnPheSerLeu 889
Qy      3859 -----GTGATTTCTGGGATACTCC 3879
Db      890 AlaGlySerGlyTyrValAsnMetGlyAlaTyrAspTyrThrLeuValGluAspAsnAsn 909
Qy      3880 GTGGGGTAT-----TCTCTTTTACTCTGTCGCAACC 3912
Db      910 AspTrpTyrLeuArgSerGlnGluValThrProProSerProAspProAspProThr 929
Qy      3913 TCATCATCAACTCCG-----CCTGCTCCAAAGTATGATGCTCGGAAAGGGTCT 3963
```

```
Db      930 ProAspProAspProThrGlnAspProAspProThr---ProAspProGluProThrPro 948
Qy      3964 ATTTTTCCTAGAGACTAGTTTGGAGATCTCAGGCGCTCAAAAAGGGGTCTATGTCAT 4023
Db      949 AlaTyrGlnProValLeuAsnAlaLysValGlyGly-----TyrLeu 962
Qy      4024 AATAATCCGGGAATTTCCGACACAGTTTTCGAGGTAAAGTAATAATAATGCTGGTGGT 4083
Db      963 AsnAsnLeuArgAlaAlaAsnGlnAlaPheMetMetGluArgArgAspHisAlaGlyGly 982
Qy      4084 GGAGGC 4089
Db      983 AspGly 984
RESULT 25
US-11-052-554A-172
; Sequence 172, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 172
; LENGTH: 1643
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
US-11-052-554A-172

Alignment Scores:
Pred. No.:      2,8e-12      Length:      1643
Score:          242.00      Matches:     282
Percent Similarity: 33.2%      Conservative: 188
Best Local Similarity: 19.9%      Mismatches: 526
Query Match:     3.1%      Indels:      419
DB:              11      Gaps:         70

US-10-701-844-1 (1-4435) x US-11-052-554A-172 (1-1643)
Qy      604 GCAGCTTTCCTTAAAGTTGTTTGGGAACCTATTAGGAGTGTCTTACTGTTTATGGGAGA 663
Db      14 AlaGlyLeuValThrAlaSerThrAlaThrIleValAlaGlyPheSerGlyValAlaMet 33
Qy      664 GCAGACTCGTTGACTTTTCGAGAACATACGGACTCTTCAAAATGGGCGAGCTCTAAGTAAT 723
Db      34 GlyAlaAlaMetGlnTyrAsnArg-----ThrThrAsnAlaAlaAla----- 47
Qy      724 AGCGTCTGATGAGTCTGCTTTTACTATTGAGGGTTTTAAAGATTTATCTTTTCCAAATGTC 783
Db      48 ThrThrPheAspGlyIle-----GlyPheAspGlnAlaAlaGlyAlaAsn--- 62
Qy      784 AATTCATTACTTTCGCTACTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACG 843
Db      63 -----IleProValAlaProAsnSerValIleThrAlaAsnAlaAsnAsnProIle 79
Qy      844 ACAACATCTACACCGCTCTAATGGTACTATTATTCT-----AAAAACAGAT 888
Db      80 ThrPheAsnThrPro---AsnGlyHisLeuAsnSerLeuPheLeuAspThrAlaAsnAsp 98
Qy      889 CTTTGTGTACTCAATAATGAG----- 909
Db      99 LeuAlaValThrIleAsnGluAspThrThrLeuGlyPheIleThrAsnIleAlaGlnGln 118
Qy      910 ---AAGTCTCTCATCTTATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAG 966
```

Db 119 AlaLysPhe---PheAsnPheThrValAlaAlaGlyLysIleLeuAsnIleThrGlyGln 137  
QY 967 AGCTTAACGGTTCAAGGAATTAAGCAAGCTTTGTGCTTCCACAGAAATACGCTCAAGCT 1026  
Db 138 GlyIleThrValGlnGluAlaSerAsnThrIleAsnAlaGlnAsnAlaLeuThrLysVal 157  
QY 1027 GATGGGGAGCTTGTCAAGTAGTCACCACTTCTCTGCTATGCTTAAC----- 1074  
Db 158 HisGlyGlyAlaAlaIleAsnAlaAsnAspLeuSerGlyLeuGlySerIleThrPheAla 177  
QY 1075 -----GAGGCTCTATT 1086  
Db 178 AlaAlaProSerValLeuGluPheAsnLeuIleAsnProThrThrGlnGluAlaProLeu 197  
QY 1087 GCCTTTGATAGCAATGTCAGAGTAAGAGGGGA---GGGATTGCTGCTGCTCAGGAT 1143  
Db 198 ThrLeuGlyAlaAsnSerLysIleValAsnGlyGlyAsnGlyThrLeuAsnIleThrAsn 217  
QY 1144 GGGCAGCAGGGAGTGCATCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCACGA 1203  
Db 218 GlyPheIleGlnValSerAspAsnThrPheAlaGlyIleLysThrIleAsnIleAspAsp 237  
QY 1204 AATACTCGGTAGAGTTGATGGGAACGTA-----GCCCGAGTA 1242  
Db 238 CysGlnGlyLeuMetPheAsnSerThrProAspAlaAlaAsnThrLeuAsnLeuGlnVal 257  
QY 1243 GAGGA-----GGGATTACTCTACGGGAACGTTGCTTTCCTG--- 1281  
Db 258 GlyGlyAsnThrIleAsnPheAsnGlyIleAspGlyThrGlyLysLeuValLeuValSer 277  
QY 1282 ---AATAATGGAAACCTGTTT----- 1302  
Db 278 LysAsnGlyAlaAlaThrGluPheAsnValThrGlyThrLeuGlyGlyAsnLeuLysGly 297  
QY 1303 -----CTCAACAATGTTGCTTCTCTCTGTTTACATGCTGCTAAGCAACCAACAGT 1353  
Db 298 IleIleGluLeuAsnThrAlaAla-----ValAlaGlyLysLeuIleSerGln 313  
QY 1354 GGCAGCGTTCTAAT-----ACGAGTAATAATTACGGAGAT----- 1389  
Db 314 GlyGlyAlaAlaAsnAlaValIleGlyThrAspAsnGlyAlaGlyArgAlaAlaGlyPhe 333  
QY 1390 -----GGAGGCTATCTTCTGTAAGAAT 1413  
Db 334 IleValSerValAspAsnGlyAsnAlaAlaThrIleSerGlyGlnValIleAlaLysAsn 353  
QY 1414 GGTGCGCAACGAGTCCAACTCTGATCAGTTTCTTGTGATGGAGGAGGAGTAGTT 1473  
Db 354 MetValIleGlnSerAlaAsnAlaGlyGlyGlnValThrPheGlu----- 368  
QY 1474 TTCTTTAGTACCAATGTAGTGTGGGAAAGGGGAGCTATTATTATGCCAAA----- 1524  
Db 369 -----HisIleValAspValGlyLeuGlyGlyThrAsnPheLysThrAlaAsp 385  
QY 1525 ---AAGCTCGGTT-----GCTAACTGGGCGCTGACAAATT----- 1560  
Db 386 SerLysValIleIleThrGluAsnSerAsnPheGlySerThrAsnPheGlyAsnLeuAsp 405  
QY 1561 -----TTAAGG 1566  
Db 406 ThrGlnIleValValProAspThrLysIleLeuLysGlyAsnPheIleGlyAspValLys 425  
QY 1567 AATATCGCTAATGATGTGGAGCGATTTATTATTAGAGAAATCTGGAGAGCTC---AGTTTA 1623  
Db 426 AsnAsnGlyAsnThrAlaGlyValIleThrPheAsnAlaAsnGlyAlaLeuValSerAla 445  
QY 1624 TCTCTGATTATGAGAGATATTATTTCGATGGGAATCTTAAAGAACAGCCCAAGAGAAAT 1693  
Db 446 SerThrAspProAsnIleAlaValThrAsnIleAsnAlaIleGluAlaGluGlyAlaGly 465  
QY 1684 GCTCCCGATGTTATGGGTAACTGTGTCCTCACAGCCATTTTCGATGGGATCGGGAGG 1743  
Db 466 ValValGluLeuSerGlyIleHisIleAlaGlu-----LeuArgLeuGlyAsnGlyGly 483

QY 1744 AAAATAACGACATTAAAGAGCTAAAGCAGGCGCATCAGATTCTTTTAATGATCCCAATCGAG 1803  
Db 484 SerIle-----PheLysLeuAlaAspGlyThrValIleAsnGlyProValAsn 499  
QY 1804 -----ATGCGAAACGGAATTAACAGCAGCGAGCTCTCCAAACTCTTAAAAATT 1854  
Db 500 GlnAsnAlaLeuMetAsnAsnAsnAlaLeuAlaGlySerIleGlnIle----- 516  
QY 1855 AACGATGTGAAGGA---TACACAGGGGATATTGTTTTTGTAAATGGAAGC---AGTACT 1908  
Db 517 ---AspGlySerAlaIleIleThrGlyAspIle-----GlyAsnGlyGlyValAsnAla 533  
QY 1909 TTGTACCAAAATGTTACGATA-----GAGCAAGGAAGATTGTTCTTCGTCAAAAGGCA 1962  
Db 534 AlaLeuGlnHisIleThrLeuAlaAsnAspAlaSerLysIleLeuAlaLeuAspGlyAla 553  
QY 1963 AAATATTACGTAATCTCTAAGTCAGACAGGTGGAGTCTGTATATGGAAGCT---GGG 2019  
Db 554 AsnIleIleGly-----AlaAsnValGlyGlyAlaIleHisPheGlnAlaAsnGly 570  
QY 2020 AGTACATCGGATTTGTAACCTCCACAAACCACACAGCCTCTCGCCGCTAATCAGTTG 2079  
Db 571 GlyThrIleLysLeuThrAsnThrGln-----AsnAsnIle 582  
QY 2080 ATCAGCTTTTCCAAATCTGCATTTGTCT-----CTTCTTCT 2115  
Db 583 ValValAsnPheAspLeuAspIleThrThrAspLysThrGlyValValAspAlaSerSer 602  
QY 2116 TTGTAGCAAAACATGAGGTACGAATCTCTCTACCAATCTCCAGCGCAAGATTCTCAT 2175  
Db 603 LeuThrAsnAsnGlnThrLeuThrIleAsnGlySer----- 614  
QY 2176 CTGCGAGTCATTGTTAGCACAACTCTGTTCTGTACAAATTAGTGGGCTATCTTTT 2235  
Db 615 -----IleGlyThrValValAlaAsnThrLysThrLeuAla 626  
QY 2236 GAGGATTGCGATGATACAGCTTATGATAGGTATGTTGGCTAGGTTCTAATCAAAAATC 2295  
Db 627 ---GlnLeuAsn-----IleGlySerSerLysThrIle 636  
QY 2296 AATGCTCTGAAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCACAGATTGACT 2355  
Db 637 -----LeuAsnAlaGlyAspVal-----AlaIleAsnGluLeuVal 648  
QY 2356 CTAGGAATGAGATGCTAAGTATGCTATCAAGGAAGCTGGAAGCTGCGTGGGATCCT 2415  
Db 649 IleGluAsnAsnGlySerValGlnLeuAsnHisAsnThrTyrLeuIleThrLysThrIle 668  
QY 2416 AATACAGCAAAATAATGCTCTTATCTCTGAAGCTTACATGGAATAAACTGGGTATAAT 2475  
Db 669 AsnAlaAlaAsnGlnGlnIleIleValAlaAlaAspProLeuAsnThrAsn----- 686  
QY 2476 CTGGGCGCTGAGCAGTAGCTCTTTGTTTCCAAATAGTTTATGGGATCATTATAGAT 2535  
Db 687 -----ThrThrLeuAlaAspGlyThrAsnLeuGlySerAlaGluAsn 700  
QY 2536 ---ATGAGTCTCGGATTCAGCAATTCAAGCA-----AGTGGGATGGGCGC 2580  
Db 701 ProLeuSerThrIleHisPheAlaThrLysAlaAlaAsnAlaAspSerIleLeuAsnVal 720  
QY 2581 TCTATTGTCGAGGATTTGGTTTCTGGAGTTTTCGAATTTCTTCTATCATCGCGCAT 2640  
Db 721 GlyLysGlyValAsnLeuIleThrAlaAsnAsnIleThr-----AsnAsp 735  
QY 2641 GCCTTAGTTCAGGATATCGGTATATATAGTGGGGTTATTCTTAGGAGCAAACTCTAC 2700  
Db 736 AlaAsnValGlySerLeuHisPheArgSerGlyThrSerIleValSerGlyThrVal 755  
QY 2701 TTTGGATCATCGATGTTTGTCTAGCATTTACCGAAGTATTTCGTAGATCTAAGATTAT 2760  
Db 756 GlyGlyGlnGln-----GlyHisLysLeuAsnAsn 765

Qy	2761	GTAGTGTGTCCTCCAAATCATCATGCTTGCATAGGATCCGTT-----TATCTATCTACC	2811
Db	766	LeuIleLeuAspAen-----GlyThrThrValIysPheLeuGlyAsp	779
Qy	2815	CAACACAGCTTTATGTGGATCCTATTTGTTCCGGAGATGCGTTATTCGTTACGTAGCTACGGG	2874
Db	780	ThrThrPheAenGlyGlyThrLysIleGluGlyLysSerIleLeuGlnIleSerAenAen	799
Qy	2875	TTTGGGAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGCATGTTCTGTTGG	2934
Db	800	ThrThrThrAspHisValGluSerAenAenAenThrGlyThrLeuGluPheValAenThr	819
Qy	2935	GATAATAAATCTGCTGGCTGGAGAGATTGGAGCGGGATTACCGATTGTGATTTACTCCATCT	2994
Db	820	Asp-----ProIleThrValThrLeuAen	827
Qy	2995	AAG-----CTCTATTGTAATGAGTTGCGTCCCTTTGTTGCAAGCTGAGTTTCTTATGCC	3048
Db	828	LysGlnGlyAlaTyrPheGlyValLeuLys-----GlnValIleIleSerGlyPro	844
Qy	3049	GATCATGAATCTTTTACAGAGGAAGGCGATCAAGTCGGGCATTTCAAGAGCGCATCTC	3108
Db	845	GlyAenIleValPheAenGluIleGly-----	853
Qy	3109	CTAAATCATCATGTTCTGTTGGAGTGAAGTTTTCATCGATGTTCTTAGTACCATCCTAAT	3168
Db	854	--AsnValGlyIleValHisGlyIleAa-----AlaAen	864
Qy	3169	AAATATAGCTTTATGGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGTACTGACG	3228
Db	865	SerIleSerPheGluAenAlaSerLeuGlyThrSerLeuPheLeuProSerGlyThrPro	884
Qy	3229	ACAACGCTCTATCCCATCAAGACACA-----TGGACAACAGATGCTTTTCATTTA	3279
Db	885	LeuAspValIleThrIleLysSerThrValGlyAenGlyThrValAspAenPheAenAla	904
Qy	3280	GCAAGACATGGAGTTGTGTTAGAGGATCTATGTATGCTTCTCTAAACAAGTAAATAGAA	3339
Db	905	Profile-----ValValValSerGly-----IleAspSerMetIleAenAenGlyGln	920
Qy	3340	GTATATGGCATGGAAGATATGATTCGAGATGCTTCTCGAGGCTATGTTTGGTAGTGCA	3399
Db	921	IleIleGly-----AspLysLysAenIleIleAlaLeuSerLeu	933
Qy	3400	GGAAAGTAGAGTCCGGTTCATAAATATGCTAGATAGTTAAGTGTAGCGATGCGCTTTT	3459
Db	934	Gly-----	934
Qy	3460	TCTTTGAGATCTACATCATTTTGTGTTTGTGTTTGTGTTCTTATTCGTATGGATTC	3519
Db	934	-----	934
Qy	3520	GCGAGCTCTCTCAAGTGTTTAAGCCCTAATGTAAACCATCCTCTTTTAAGGAGACGATGT	3579
Db	935	-----SerAspAenSerIleThrValAenAlaAenThrLeuTyrSerGlyIleAenThr	952
Qy	3580	TACTTGATGGAGACTCGGCTTTTGTC-----AAT	3609
Db	953	ThrLysAenAenGlnGlyThrValThrLeuSerGlyGlyMetProAenAenProGlyThr	972
Qy	3610	GTCTATGCA---GGAGCTGAAGAGGTTCGATTATCTCAGCTAATGCGGACAATTTAAACG	3666
Db	973	IleTyrGlyLeuGlyLeuGluAenGly-----SerProLysLeuLysGlnValThr	989
Qy	3667	ATTACCGGACAAACCATCATATCATTTACAGATTCTCAAGGGCCAGTTCTTCTCAAAAT	3726
Db	990	PheThrThrAspTyrAenAenLeu-----GlySerIleIleAlaAen	1003
Qy	3727	TATGCTTCTATTTTCAGCAGGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTTC	3786
Db	1004	AsnValThrIleAenAspTyrValThrLeuThrThrGlyGlyIleAlaGlyThrAspPhe	1023
Qy	3787	TCGAAAAATGTTTCTTCGGAGAAAAGGGAATGATCTCCGGGAAAAACCGTGAGTATTTCC	3846

Db	1024	AspAlaValyleThrLeuGly				-----SerValAsn	1033
QY	3847	GGAGCGCGAAGTGATTTTCGGATAACTCGTGGGTATTCTCTTTATCTACTGTG				-----SerValAsn	3906
Db	1034	GlyAsnAlaAsnValArgPheValAspSerThrPheSer				-----AspProArgSerMetIle	1052
QY	3907	CCAACTTCATCACTCACTCGCCTGCTCCACACAGTTAGTAGTCTCGGAAGGGTCTATT				-----Ser	3966
Db	1053	-----ValAlaThrGlnAlaAsnVal				-----Ser	1062
QY	3967	TTTTCTGTAGAGACTAGTTTGGAGATCTAGCGCTCAAAAAGGGGTCAATGTTTCGATAAT				-----Ser	4026
Db	1063	ThrTyrLeuGlyAsnAlaLeuVal				-----Ser	1071
QY	4027	AATCGCGGAATTCGGAACAGTT				-----Ser	4071
Db	1072	AsnIleGlySerLeuAspThrProValAlaSerValArgPheThrGly				-----Asn	1088
QY	4072	AATGCTGCTGGAGGACGTGGTTCGCTACACCATCAAGTACGACTTTTACAGTTAA				-----Asn	4131
Db	1089	AspSerGlyAlaGlyLeuGlnGlyAsnIleTyrSerGlnAsnIleAspPheGlyThrTyr				-----Ser	1108
QY	4132	AAACTGTAAAGGAAAGTTCTTTTCACAGATAACGTAGTCCCTCTTGGCGGAGCGCGAGTGGT				-----Ser	4191
Db	1109	AsnLeu-----ThrIleLeuAsn-SerAsnValIleLeuGlyGlyThrAla				-----Ser	1125
QY	4192	TTATAAGGC--ATTCTGCTTTTCAAAGACATGAAGGAGGATATCTTCCGAGGAA				-----Ser	4248
Db	1125	alleAsnGlyGluIleAspLeuLeuThrAsnAsnLeu-----IlePheAlaAsnGlyTh				-----Ser	1143
QY	4249	CACAGCATACGATGATTAAAGGATCTTGCTGCTACTAATCAGGATCAGAATACGGAGAC				-----Ser	4308
Db	1143	rSerThrTyrGlyAspAsnThrSerIleSerThr-----LeuAsnValSerSe				-----Ser	1160
QY	4309	AGGAGCGGTGGAGGAGTATTTGCTCTCAGAGTATCTGTGA				-----Ser	4351
Db	1160	rGlyAsnIleGlyGlnValValIleAlaGluAspAlaGlnVal				-----Ser	1174

RESULT 26

US-11-052-554A-281

; Sequence 281, Application US/11052554A

; Publication No. US2005028866A1

; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.

; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN

; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A

; CURRENT FILING DATE: 2005-02-07

; PRIOR APPLICATION NUMBER: US 60/589,227

; PRIOR FILING DATE: 2004-07-20

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004

; PRIOR FILING DATE: 2004-02-06

; NUMBER OF SEQ ID NOS: 763

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 281

; LENGTH: 5291

; TYPE: PRT

; ORGANISM: Escherichia coli 0157:H7

US-11-052-554A-281

Alignment Scores:			
Pred. No.:	1.76e-11	Length:	5291
Score:	235.50	Matches:	287
Percent Similarity:	34.2%	Conservative:	193
Best Local Similarity:	20.5%	Mismatches:	537
Query Match:	3.0%	Indels:	386
DB:	11	Gaps:	63

US-10-701-844-1 (1-4435) x US-11-052-554A-281 (1-5291)

QY 496 GGGGAGACGTTAACTGTA-----TCATTTCCCTACTACTGTATAGGAGAT 540









Db 990 GlnTyrThrLeuGlyArgSerLysAspGluPheGlnAlaLeuAlaArgAlaGluAspLeu 1009  
Qy 1885 GTTTTGTCTTAATGGAAGCAGTACTTTGTAC----- 1914  
Db 1010 GlnValAlaGlyGlyThrAlaIleValTyrAlaGlyThrLeuAlaAspAlaSerValSer 1029  
Qy 1915 -----CAAATGTTACATAGAG----- 1932  
Db 1030 GlyAlaThrGlySerLeuSerLeuMetThrProArgAspAsnValThrProValLysLeu 1049  
Qy 1933 CAAGGAGGATTGTTCTTCGCGAAAGCAAAATATCAGTG---AATTCTCTAAGTCAG 1989  
Db 1050 GluGlyAlaIleArgIleThrAspSerAlaThrLeuThrIleGlyAsnGlyValAspThr 1069  
Qy 1990 ACAGTGGGAGTCTGTATATGTAAGCT---GGGAGTACATGGGATTTTGTAACTCCACA 2046  
Db 1070 ThrLeuAlaAspLeuThrAlaAlaSerArgGlySerValTyr----- 1083  
Qy 2047 CCACCACACAGCTCTCGCGCTTAATCAGTTGATCAGCTTTCATCTGCATTTGTCT 2106  
Db 1084 -----LeuAsnSerAsnAsnSerCysAlaGlyThrSerAsnCysGluTyrArg 1099  
Qy 2107 CTTTCTTCTTGTAGCAAC-----AATGCAGTTACGAATCTCTCTACCAATCTCCA 2160  
Db 1100 ValAsnSerLeuLeuLeuAsnAspGlyAsnValTyrLeuSerAlaGlnThrAlaAlaPro 1119  
Qy 2161 GCGCAAGATTCTCATCTCGCATGCTGTCAGACACACTGCTGTTCTGTTCAATTAAGT 2220  
Db 1120 AlaThrThrAsn-----GlyIleTyrAsnThrLeuThrThrAsnGluLeuSerGlySer 1137  
Qy 2221 GGGCCTATCTTTTGTAGGATTGGATGATCAGCTTATGATAGTATGATGATGGCTAGT 2280  
Db 1138 GlyAsnPheTyrLeuHis-----ThrAsnValAlaGlySerArgGlyAspGlnLeuVal 1155  
Qy 2281 TCTAATCAA-----AAATCAATGCTCCTGAAATTCACATTAAGGACTAAGCCC 2328  
Db 1156 ValAsnAsnAsnAlaThrGlyAsnPheIlePheValGlnAspThrGlyValSerPro 1175  
Qy 2329 CCAGCTAAT-----GCCCCATCAGATTTTGACTCTA 2358  
Db 1176 GlnSerAspAlaMetThrLeuValLysThrGlyGlyGlyAspAlaSerPheSerLeu 1195  
Qy 2359 GCGAAT-----CAGATGCTAAGTATGCTAT----- 2385  
Db 1196 GlyAsnThrGlyGlyPheValAspLeuGlyThrTyrGluTyrValLeuLysSerAspGly 1215  
Qy 2386 CAAGGAAGCTGAAGCTTCGCTGGAT-----CCTAAT---ACAGCAATAATGCTCT 2436  
Db 1216 AsnSerAsnTrpAsnLeuThrAsnAspValLysProAsnProAspProAsnPro 1235  
Qy 2437 TATACTCTGAAAGCTACATGGAGTAAACTGGGTATTAATCTCGGCTGAG----- 2487  
Db 1236 AsnProAsnProLysProAspProLysProAspProLysProAspProLysProAspPro 1255  
Qy 2487 ----- 2487  
Db 1256 ThrProGluProThrProThrProValProGluLysArgIleThrProSerThrAlaAla 1275  
Qy 2488 -----CGAGTAGCTTCTTTGTTTCA----- 2508  
Db 1276 ValLeuAsnMetAlaAlaThrLeuProLeuValPheAspAlaGluLeuAsnSerIleArg 1295  
Qy 2509 -----AATAGTTTATGGGATCCATTTTA 2532  
Db 1296 GluArgLeuAsnIleMetLysAlaSerProHisAsnAsnAsnValTrpGlyAlaThrTyr 1315  
Qy 2533 GATATACGATCT-----GGCATTTACGAATTTCAAGCAAGTGTGGATGGG 2577  
Db 1316 AsnThrArgAsnAsnValThrThrAspAlaGlyAlaGlyPheGluGlnThrLeuThrGly 1335  
Qy 2578 -----CGCTCTATTGTCGAGGATTATGGGTTTCTGAGTT 2613  
Db 1336 MetThrValGlyIleAspSerProAsnAspIleProGluGlyIleAlaThrLeuGlyAla 1355

Qy 2614 TCGAATTTCTTCTATCATGACCGCGATGCTTTAGTTCAGGATATCGTATATTAGTGGG 2673  
Db 1356 PheMetGlyTyrSerHisSerHisIleGlyPheAspArgGlyGlyHisGlySerValGly 1375  
Qy 2674 GGTATTCTCTAGGAGCA---AATCTCTACTTTGGATCATCGATGTTTGGTCTAGCATTT 2730  
Db 1376 SerTyrSerLeuGlyGlyTyrAlaSerTrpGluHisGluSerGlyPheTyrLeuAspGly 1395  
Qy 2731 ACCGAAGTATTGTTGATGATCTAAAGATTATGATGCTGTCCTCAATCATCATGCTTGC 2790  
Db 1396 ValValLysLeuAsnAsnArgPheGluSerAsnValAlaGlyLys----- 1409  
Qy 2791 ATAGATCCGTTTATCTATCTATCTACCAACAGCTTTATGTGTGATCTTATTTCTGGAGAT 2850  
Db 1410 -----MetSerSerGlyGlyAlaAlaAsnGlySerTyr----- 1420  
Qy 2851 GCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTT 2910  
Db 1421 -----HisSerAsnGlyLeuGlyGly---HisIleGluThrGlyMetArgPhe 1435  
Qy 2911 GCAGAGGAGCGATGTTCTGTTGGAT 2937  
Db 1436 -----ThrAspGlyAsnTrpAsn 1441  
RESULT 28  
US-11-052-554A-171  
; Sequence 171, Application US/11052554A  
; Publication No. US20050288866A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 171  
; TYPE: PRT  
; ORGANISM: Rickettsia prowazekii  
US-11-052-554A-171  
Alignment Scores:  
Pred. No.: 5,54e-10 Length: 2340  
Score: 216.50 Matches: 276  
Percent Similarity: 31.9% Conservative: 191  
Best Local Similarity: 18.9% Mismatches: 569  
Query Match: 2.7% Indels: 426  
DB: Gaps: 63  
US-10-701-844-1 (1-4435) x US-11-052-554A-171 (1-2340)  
Qy 496 GGGGAGAGCTTAACGTGATCATTT---CCCTATCTGTTATAGGATCCGAGTGGGACT 552  
Db 105 GlyGluAspLeuAsnThrAsnPheGlyProLeuLysPheIleSerAsnAsnValThrSer 124  
Qy 553 ACTGTTTCTCTGCGAGGAGGTTAACTATAAAAAATCTTGAC-----AATTTCTATTGCA 606  
Db 125 IleIleThrGlyValGlyThrLysThrPheSerAsnIleAspPheAlaGlyLysAsnAla 144  
Qy 607 GCTTTGCCCTTAAGTGTGTTTGGGAACCTTATTAGGAGTTT-----ACTGTTTATGGG 660  
Db 145 ThrLeuGlnIleAsnLysAspLeuAsnIleThrThrLysIleAspAsnThrValAlaGly 164  
Qy 661 AGAGGACACTCGTTCACCTTTCGAGAAACATACGACTTCTACAAATGGGCGAGCTCTAACT 720  
Db 165 AsnAsnGlySerIleThrPheGluGlySerGlyIleIleSerAsnHisIleGlyTyrThr 184

```
QY 721 AATACGCGTGCGTGGAGCTG----- 741
Db 185 AenSerLeuLeu---GlyIleAenValGlyAenGlyGluAlaLysIleTyrAlaProGlu 203
QY 742 -----TTTACTATTGAGGTTTAAAGAAATTATCTCTTCCAAATTGCAATTCAATTA 792
Db 204 AlaAenAenIleThrIleAenAla---LysAenIleAenLeuThrHisAenAenSerIle 222
QY 793 CTTGCGGTACTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGCAACATCT 852
Db 223 LeuThrLeuCysAenGlyAenIleThrThrLeuLysGlyAenIleAenAenThrThrGlu 242
QY 853 ACACGCTCTAATGCTACTATTATTTCTTAAACAGACTCTTTTCTACTCAATATGAGAAG 912
Db 243 IleAenGlyGlnGlyIleLeuAenLeuAlaTyrAenLeu----- 255
QY 913 TTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGACTTGA 972
Db 256 -----GlySerSerIleIleThrGlyAenIleGlyAenIleGlySerLeuAen--- 272
QY 973 ACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAAAAATFACTGCTCAAGCTGATGGG 1032
Db 273 ThrIleAenValLeuLeuGlySerAlaThrPheAenSerThrIleLeuLysAlaThrAen 292
QY 1033 GGAGCTTGCTCAAGTAGTACCAGTTCTCTGCTATGGCTAAACGAGGCTCCTATTGCCCTT 1092
Db 293 -----IleAenLeuLysHisAenThrSerThrLeuAenLeuAenAenIleIleVal 310
QY 1093 GTAGCGAATGTTGCAGGA-----GTAAGAGGGGGA 1122
Db 311 IleGlyAenIleLysGlyAenAenAenLysAenAenIleLeuAenPheLysValHisGlyThr 330
QY 1123 GGGATTGCTGCTGTTCCAGGAGGCGACGAGGAGTCTCATCTACTTCAACAGAGAT 1182
Db 331 AenLeu-----AspAenGluMetIleIleProAlaProGlnLysThrHisGly 346
QY 1183 CAGTAGTAGTATGTTTCCAGAAATACTGCGGTAGAGTTGTAGTGGGACGTAGCCCGAGTA 1242
Db 347 ---ThrLeuAenPheLysGlyAenAlaThrLeu-----AenGlyAenIleAenAenLeu 363
QY 1243 -----CGAGCA----- 1248
Db 364 AenIleLeuLysPheSerGlyGlyHisGlyLysThrLeuAenLeuGlnGlyAenThrLys 383
QY 1249 -----GGG 1251
Db 384 ValAenAenLeuValPheAlaAenSerValLeuAenSerGlyThrIleSerValAenGly 403
QY 1252 ATTTACTCTACGGGAACGTTGCTTTCCTGAAT-----AATGGAAAAACCTTGTTT 1302
Db 404 LeuLeuAenThrAenCysValThrPheAenAenSerAenValAenGlyGlyThrLeuIle 423
QY 1303 CTC-----AACATGTTGCTCTCTCTGTTTACATTCCTGCTCAAGCAACCAACAGT 1353
Db 424 IleAenAlaLysAenThrIleSerAlaLysLeuLeuAenAenAlaThrLysAlaLysIleGln 443
QY 1354 GGCAGCGCTTCTAATACGAGTAATAATTACGGAGATGGAGGA----- 1395
Db 444 IleAenAlaAenLeuThrMetAenHisProSerAlaGlyAenPheSerAenIleArgIle 463
QY 1396 -----GCTATCTCTGTAAGAATGTCGCGACGAGATCAATAAC 1437
Db 464 AlaAenAenThrIleTyrThrIleAenAlaLysAenGlyAenValAenLeuAenAenAen 483
QY 1438 TCTGGATCATCTTCTTGTATGAGAGAGGA-----GTAGTTTCTTTAGTACATGTA 1491
Db 484 AenAlaLysIleIlePheGluGlyAlaAenSerMetLeuAlaLeuIleAenThrGlyVal 503
QY 1492 GCTGCTGGGAAGGGGAGCTATTATTATGCCAAA-----AAGCTC 1530
Db 504 ThrAlaAenArgThrPheThrIleTyrAenAenLeuAenGlnSerGlyAenAenGluTyr 523
```

---

```
QY 1531 TCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATATGGTGGAGCG 1590
Db 524 GlyIleValLysIleGluAlaIleLysLysValIleThrIleAlaAenGlnSerGlyPro 543
QY 1591 ATTTATTAGGAGAACTGTGGAGAGCTCAGTTTACT-----GCTGATTATGGA 1638
Db 544 TyrThrIleGlyGlnAenAenThrHisArgLeuLysGluLeuIleValGluGlyAlaGly 563
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAT 1698
Db 564 AspIleIleIleAenAenAen----- 569
QY 1699 GCGGTAACTGTGCTCTCAAAAGCCATTTTCGATGGGATCGGAGGGAATAACGACATTA 1758
Db 570 -----ThrIlePheThrLysLeuLeuSerIleAenSerThrGlyGlnIleThr----- 585
QY 1759 AGAGCTAAAGCAGGCGCATCAGATTCTTTTAAATGATCCATCGAGATGGCAACGGAAT 1818
Db 586 -----PheAenArgThrLeuAenPheLeuAenGlyAlaGlyGly 596
QY 1819 AACCCAGCCAGCGCAGTCTTCCAAACTCTTAAAAATTAACGATGCTGAAGGATACACAGGG 1878
Db 597 AenIleAlaPheGlyLysHisGlyThrLeuValValAen-----GlyValThrGly 613
QY 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTTACGATAGCAGCAAG 1938
Db 614 SerIleThrThrSerGluAenAenGlnGlyIle----- 624
QY 1939 AGGATTGTTCTCGTGAAGCAAAATTAATCAGTAATTTCTAAGTCAGACAGGTGGG 1998
Db 625 -----LeuThrIleAenSerGlyAenIleThrGlyVal 635
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACCAACCAACAG 2058
Db 636 IleGlyThrAenGluLeuGlyLeuLysLeuValAenIleGlyAlaAenPro----- 652
QY 2059 CCTCCTGCGCTAACTCAGTTGATCAGCTTTCCTCAATCTGCATTTGTCTCTTCTCTTTG 2118
Db 653 -----ValThrCysSer-----AlaAenVal 659
QY 2119 TTAGCAAAACAATGCTATGCTCTCTCTCAATCTCTCCAGCGCAAGATTTCTCATCTC 2178
Db 660 PheAlaSerValAlaLeuThrAenProSerSerValLeuIleLeuAlaAen----- 676
QY 2179 GCAGTCATTGCTAGCACAACCTGCTGTTCTGTGTACAAATTTAGTGGGCTATCTTTTGTAG 2238
Db 677 -----GlyValThrLeuThrGlyGluValThrThr----- 686
QY 2239 GATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTAATCAAAAATCAAT 2298
Db 687 -----HisAenAenThrLysGly 692
QY 2299 GTCCTGAAATTA-----CAGTTAGGGACTAAGCCCCCAGCT 2334
Db 693 ValLeuSerLeuGlyThrGlySerAenIleThrGlyGlnIleGlyThrAenSerAlaAla 712
QY 2335 AATGCCCATCAGATTGCTCTAGG-----AATGAGATGCTTAAGTATGCTATCAAGGA 2391
Db 713 LeuGluLysIleAenIleGlyAlaGlyAlaSerAenIleAenAenAenAenIleThrAlaGly 732
QY 2392 AGCTGGAAGCTTGGTGGGATCTCTAAT-----ACAGCAAAATAATGGTCTTATCTACTCTG 2445
Db 733 SerThrValLeuThrAenPheGlnThrSerGluLeuThrLeuAenAenAenAenValValAen 752
QY 2446 AAGCTTACATGGAATAAACTGGGTATATCTCGGCTGAGGAGTAGCTTCTTTGGTT 2505
Db 753 SerAenIleIleThrThrAlaGlyAenAenSerGly-----LysLeuIle 767
QY 2506 CCAATAGTTTATGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCA 2565
Db 768 PheThrGlyAenGlyGlyIleThrGlyAenIleGlyAlaAenGlyAlaLeuGlnGlu 787
QY 2566 AGTGTGATGGGCGCTCTTATTCTCGAGGATTTATGGTTTCTCGAGTTTCGAATTTCTTC 2625
```



;; PRIOR APPLICATION NUMBER: US 60/589,227  
;; PRIOR FILING DATE: 2004-07-20  
;; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
;; PRIOR FILING DATE: 2004-02-06  
;; NUMBER OF SEQ ID NOS: 763  
;; SOFTWARE: Patent in version 3.3  
;; SEQ ID NO 283  
;; LENGTH: 1461  
;; TYPE: PRN  
;; ORGANISM: Escherichia coli 0157:H7  
US-11-052-554A-283

Alignment Scores:  
Pred. No.: 9,99e-10 Length: 1461  
Score: 212.50 Matches: 295  
Percent Similarity: 33.7% Conservative: 200  
Best Local Similarity: 20.1% Mismatches: 540  
Query Match: 2.7% Indels: 432  
DB: 11 Gaps: 75

US-10-701-844-1 (1-4435) x US-11-052-554A-283 (1-1461)

```
QY 439 TCTTTAAATGGGGGGATATGCA---GCAGAAATCATGGTTCCTCAGGAATTTACGAT 495
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 ThrPheSerGlySerGlyGlnAlaGlyAlaThrIleGlnIleLeuSer-----Asn 220
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 GGGGAGAGCTTA-----ACTGTATCATTTCCC 522
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 GlyAsnThrIleAlaSerThrGlnValAspAsnAsnGlyHisTrpSerValSerLeuPro 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 TATACGTATTATAGAGATCCGAGT-----GGGACTACTGTGTTTT 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ThrGlnSerAlaGlyGluHisThrTrpSerValValGlnIleValGlySerThrIleThr 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 TCTGAGAGAGTTAACTATAAAATCTTGACAAATCTTAATTCGAGCTTTCCTTAAGT 621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 AspAlaGlySerIleThrLeu---ThrIleAspAsnSerGlnAlaSerValGlnValAla 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 622 TGTTTTGGG-----AACTTATTA 639
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 ThrThrAlaGlyAspAsnIleIleAsnAlaSerGluGlnAlaAlaGlyPheThrLeuSer 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 GGGAGTTTACTGTTTGGGAGAGACACTCGTTGACTTTCGAGAACATACGAGCTTCT 699
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 GlyThrSerSerHisLeuAlaGlnGlyThrGluLeuThr-----ValThr 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 ACAAAATGGGGAGCTCTAAGTAATAGC---GCTGCTGATGAGCTGTTTACTATTGAGGT 756
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 LeuAsnGlyLysThrTyThrThrSerValGlyAlaAsnGlyAlaTrpSerValGln--- 333
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 TTTAAAGAAATTATCTTTTCCAAATTCGAATTCATTACTTGGCGTACTGCTGTCMACG 816
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 -----ValProThrAlaAspAlaGlnAla 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 ACTAATAAGGTTAGCCAGACTCCGACGACAAATCATCCACCGTCTAATGCTACTATTAT 876
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 342 LeuGlyGluGlyAsnGlnAlaValLeuValSerGlyLysAspAlaThrGlyAsnThrVal 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 TCTAAAACAGACTCTTTGTTACTCAATAATGAGAAGTCTCTCAATCTTATAGTAATTTAGTC 936
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 ThrGlyAlaGlnLeuLeuThrValAspThrGlnProProThrLeuAlaIleAsnThrIle 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 TCTGGAGAT-----GGGGAGCTATAGATGCTTAAGAGCTTAACGGTTACAGGAAT 987
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 AlaGlnAspAsnIleIleSerAlaAlaGluHisAsnValAlaLeuValLeuSerGlyThr 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 ACCAAGCTTTGTCTCTTCCAGAAATATCTGCTCAAGCTGATGGG-----1032
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 402 SerAspAlaGluAlaGlyGlnThrValThrLeuThrValAsnGlyLysSerHisThrAla 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1033 -----GGAGCTTGTCAAGTA-----GTCAACAGATTTCCTGCTATG 1068
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 ThrValGlySerAspGlyThrTrpGlnValThrLeuProAlaThrGluValGlnAlaLeu 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 1069 GCT-----AACGAGGCTCTTATTGCCCTTTGTAGCAATGTTGCAGGATGAAGAGGGGA 1122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 AlaGluGlyAsnTyAlaValAsnAlaSerValSerAspArgAlaGly-----457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1123 GGGATTGCTGCTGTTCAGGATGGCAGCAGGAGTGTCTATCTACTTCCACAGAAGAT 1182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 -----AsnThrThrSerHisSerAlaAsnPheThrValAspThrSerAla 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1183 CCAGTAGTAAGTATCTCCAGAAATATCTGCGGTAGAGTTTGTATGGGAACGTAGCCCGAGTA 1242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 ProValValSerVal-----AsnThrValAlaGlyAspAsp-----484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1243 GGAGGAGGATTTACTCTCTACGGGAACGTTCTTCTGAATAATATGGAATAAACCTTGT 1302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 485 -----IleLeuAsnAsnAlaGluGln-----491
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1303 CTCACAATGTTGCTTCTCTGTTTACATTCTGCTAAGCAACCAACCAAGTGGCAGGCT 1362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 -----AlaValAlaGlnIleIleSerGlyGlnVal 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1363 TCTAATACGAGTAATAATTACGAGATGGAGGAGCTATCTTCTGTAAGAATGGTCGCA 1422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 SerGlyAlaSer-----ProGlyAspThrValThrValValLeuGlyThrHis 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1423 -----GCAGGATCCAATAACTCTCGATCAGTTTCTCTTGTAT-----1458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 518 ValLeuThrGlyIleValLeuAlaAspGlySerTrpAsnValAlaLeuAspProAlaVal 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1459 -----GGAGAGGAGTAGTTTCTTTTAGTAGCAATGCTAGCTCTGGG 1500
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 ThrArgThrLeuAspArgGlyAlaAsnThrIlePheValThrValThrAspAlaAlaGly 557
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1501 AAAGGGGAGCTATTTATGCCAAAAGCTCTCGGTTGCTAACTGGCGCTGTACAATTT 1560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 AsnThrGlyAlaAla---SerArgAlaIleThrLeuValGlyValSerProLeuIleThr 576
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1561 TTAAGGAATATCGCTTAATGATGGT-----1584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 IleAsnThrValSerGlyAspAspIleIleSerGlyAlaGluLysGlyAlaProLeuThr 596
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1585 -----GGAGCGATTTATTAGGAGAACTCGAGAG-----CTCAGTTTA 1623
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 597 LeuThrGlySerThrGlnGlnAlaGluThrGlyGlnThrValThrValThrLeuAlaGly 616
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1624 TCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACAA-----1671
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 GlnSerPheThrThrThrValGlnAlaAspGlySerTrpSerLeuThrValProAlaAla 636
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1672 GCCAAAGAGAAATGCTGCGGAT---GTTAATGGCGTAACCTGTCTCTCAAGCCATTCG 1728
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 AlaMetGlyAsnLeuProAspGlyAlaValAlaIleThrAlaSerValThrAspLeuSer 656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1729 ATGGATCGGGAGGAGAAATAACGACATTAAAGAGCTAAAGCAGGCGGATCAGATCTCTTT 1788
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 657 GlyAsnThrGlyAsnThrSerArgThrIleThrValAspSerGlnAlaProAlaLeuSer 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1789 AATGATCCCATCGATGCGCAAC-----GGAAATACCCAGCCA 1827
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 IleAspProLeuThrAlaAspAsnIleIleAsnAlaAlaGluSerGlyGlnAspLeuPro 696
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1828 -----GCGCAGCTCTTCCAAACTTCTAAAAAATAAC---GATGTGTA 1866
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 697 IleThrGlyThrThrAspAlaGlnProGlyGlnThrValThrValThrLeuAsnGlyGln 716
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1867 GGATACACAGCGGGATATT-----1884
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 717 ThrTyThrGlnGlyValValGlnProAspGlyThrTrpSerValThrValProAlaAlaAsn 736
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1885 -----GTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAAGCA 1938
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 ValGlyAlaLeuAlaAspGlyAsnAlaThrValThrAlaSerValAsnAspValAlaGly 756
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY 1939 -----AGGATTGTTCTT----- 1950  
Db 757 AsnProSerSerValSerArgValAlaLeuValAspAlaThrProProValValThrIle 776  
QY 1951 -----CGTGAAGGCGAAATATCATCGTGAAT 1977  
Db 777 AsnProValAlaThrAspAsnValIleAsnThrProGluHisAlaGlnAlaGlnIle 796  
QY 1978 TCCTTAAGTCAGACAGCTGGAGCTGTATATGGAAGCTGGAGTACATGGATTTTGTA 2037  
Db 797 SerGlyThrValThrGlyAla-----GlnAlaGly-----AspIleVal 809  
QY 2038 ACTCCACAACCAACCAACAGCTCTCGCGCTAATCATGTTGATCAGCTTTCACAAATCTG 2037  
Db 810 Thr-----ValThrLeuAsnAsnVal 816  
QY 2098 CATTTGTCTCTTCTTCTTTAGCAACAATGAGTGTAGCAATCTCTCCACCAATCCT 2157  
Db 817 AspTyrThrThrValValAspGlySerGlyAsnTyrSerLeuGlyValProAlaSerVal 836  
QY 2158 CCA-----GCGCAAGATTCTCATCTCGAGTCATT-----GGT 2190  
Db 837 ValSerGlyLeuAlaAspGlySerTyrProValSerValSerValThrAspPlyAlaGly 856  
QY 2191 AGCAAACTGCTGGTCTGTGTACAAAT-----AGTGGCGCTATCTTTTGGAG 2238  
Db 857 AsnThrGlySerGlnSerLeuThrValThrValAsnThrAlaAlaProLeuIle----- 874  
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGTAGTGTCTTAATCAAAAATCAAT 2298  
Db 875 GlyIleAsnSerIleAlaGlyAsp-----AspValIleAsnAlaSerGluLysGlyAla 892  
QY 2299 GTCCGAAATACAGTTAGGAGCTAAGCCCGCTAATGCCCCATCAGATTCGACTCTA 2358  
Db 893 AspLeuGlnIleThrGlyThrSerAspGlnProValAsnThrAlaIleThrValThrLeu 912  
QY 2359 GGG-----AATGAGATGCTAAGTATGGCTATCAAGGAAGCTGGAAGCTT----- 2403  
Db 913 AsnGlyGlnAsnTyrThrThrThrAspAlaSerGlyAsnTyrSerValThrValPro 932  
QY 2404 CGTGGGATCTCTAATACAGCAAAATATGGTCTTATCTCTGAAGCTACATGGACTAAA 2463  
Db 933 AlaSerAlaValThrAlaLeuGlyGlnAlaAsnTyrThrValThrAlaAlaValThrSer 952  
QY 2464 ACTGGGTATATCTCGGCGCTGAGGAGTAGTCTTCTGGTCCAAATGTTTATGGGA 2523  
Db 953 AspIleGlyAsnSerAlaThrAlaSerHisAsnValLeuValAspSerAlaLeuProGly 972  
QY 2524 TCCATTTTAGAT-----ATACGATCTGCGCATTCAGCAATTCAA 2562  
Db 973 ValThrIleAsnProValAlaThrAspAspIleIleAsnAlaGluAlaGlyValAla 992  
QY 2563 GCAAGTGTGATGGCGCTCTTATTGTGAGGATATGGGTTCTCGGAGTTTCGAATTC 2622  
Db 993 GlnThrIleSerGlyGln-----ValThrGlyAla----- 1002  
QY 2623 TTCTATCATGACCGGATGCTTTAGGTAGGATATCGGTATATAGTGGGGTTATCC 2682  
Db 1003 -----GluAspGlyAspThrVal-----ThrIleThr 1011  
QY 2683 TTAGGAGCAAACTCTCTACTTTGGATCATCGATGTTTGGCTAGCATTTTACCGAAGTATT 2742  
Db 1012 LeuGlyGlyAsnThrTyr-----ThrAlaThrVal 1021  
QY 2743 GGTAGATCTAAAGATTATGATAGTGTGCTTCCAATCATCATGCTTGTGATAGATCCGTT 2802  
Db 1022 GlySerAsnLeuThrTyrSerVal-----AspVal 1031  
QY 2803 TATCTATCTACCCCAAGCTTTATGCTGATCTTATTTGTTCTGGAGATGGTTTATCCGT 2862  
Db 1032 ProAlaAlaAspIleGlnAlaLeu-----GlyAsn----- 1047  
QY 2863 GCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATATCATTTGCGAGGAGAGC 2922

Db 1048 AlaSer-----ValThrAsnGlnAsnGlyAsnThrGly-----SerGlyThrArg 1062  
QY 2923 GATGTTCTGTTGGGATAATAAC-----TGTCTGGCTGGAGAG----- 2958  
Db 1063 AspIleThrIleAspAlaAsnLeuProGlyLeuArgValAspThrValAlaGlyAspAsp 1082  
QY 2959 -----ATTGAGCGGATTAACCATGTGTGATTACTCCATCTAAGCTCTATTG 3006  
Db 1083 ValValAsnIleIleGluHisGlyGlnAlaLeuValValThrGlySerSerGlyLeu 1102  
QY 3007 AATGAGTTGGTCTCTTC-----GTCAA 3030  
Db 1103 AlaGluSerThrProLeuThrValThrIleAsnAsnValGluTyrThrThrAlaValGln 1122  
QY 3031 GCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGGCA 3090  
Db 1123 AlaAspGlySerTyr-----SerValGlyValThrAlaAlaGlnValSerAla 1138  
QY 3091 TTCAAGCGGACATCTCTAAATCTATCAGTT-----CCTGTTGGAGTGAAG 3138  
Db 1139 TrpProAlaGly-----ThrValAsnIleAlaValSerGlyGluSerSerAlaGly----- 1155  
QY 3139 TTTGATCGATGTTCTTAGTACACATCTTAATAATATAGCTTTATGGCGCTTATATCTGT 3198  
Db 1156 ---AsnSerValSerIleThrHisProValThrValAspLeuThrProAlaAlaIle--- 1173  
QY 3199 GATGCTTATCGCACATCTCTGTGACTGAGACAACGCTCTATCCCATCAAGAGACATGG 3258  
Db 1174 ---ThrIleAsnThrIleAla----- 1179  
QY 3259 ACAACAGATGCTTCTTATTTAGCAAGACATGAGATTGTGTTAGAGGATCTATGATGCT 3318  
Db 1180 ThrAspAspValIleAsnAlaAlaGluLysGlyAlaAspLeu-----ThrLeuSer 1196  
QY 3319 TCTCTAACAGTATATAGAA-----GTATATGGCCATCGAAGA 3357  
Db 1197 GlyThrThrThrAsnValGluProGlyGlnThrValThrValThrPheGlyGlyLeu 1216  
QY 3358 TATGAGTATCGAGATGCTCTCTCGAGGCTATGTTTGTAGTCGAGGAAGTAGAGTCGGTTC 3417  
Db 1217 TyrThrAlaSerValAlaSerAspGly----- 1225  
QY 3418 TAAATAATTTGTTAGATAGTTAAGTGTTCGAGTGCCTTTTCTTTGAGATCTACATCA 3477  
Db 1226 -----SerTyrThrAlaThrValProAlaAlaAspLeuAlaSerLeuProGluGlySer 1243  
QY 3478 TTTTGTGTTTGTAGCTTGTGTTGTTCTTATTCGTATGATTTCGCGAGCTCTCTCAAGTG 3537  
Db 1244 AlaSerAlaLeuAlaSerValSerAsnIleAsnGlyAsnSerAlaSerAla----- 1260  
QY 3538 TTAACGCTAATGTAAACCACTCTTTTAAAGGAGACGATGTTTACTTGAATGAGACTGC 3597  
Db 1261 ---ValHisAsnTyrSerValAspSerSerAlaProThrIleIleIleAsn---ThrVal 1278  
QY 3598 GCTTTTGTCAATGCTATCTCAGGAGCTGAAGAAGTTCGATTCATCTACGCTAATGGGAC 3657  
Db 1279 AlaSerAspAsnIleValAsnAlaSerGluAlaAspAlaGlyValThrValSerGlySer 1298  
QY 3658 AATTTAAGATTACCGGACAAACCATATCATTTATCTTACAGATCTCAAGGGCCAGTT 3717  
Db 1299 ThrThrAlaGluAlaGlyGlnIleValThrIleThrLeu-----AsnSerProThr 1315  
QY 3718 CTTCAAAATTAAT---GCCTTCATTTTACGAGGAGACACACTTACTCTGAGAGATTTTTCG 3774  
Db 1316 ValGlnThrTyrGlnAlaThrValGlnAlaAspGlySer----- 1328  
QY 3775 AGTCTGATGTTCTCGAAAAATGTTTCTTTCGCGAGAAAAGGAATGATCTCCGGGAAAAACC 3834  
Db 1329 -----TipSerIleAsnIleProAlaAlaAspLeuGluAlaLeuThr----- 1342  
QY 3835 GTGAGTATTTCCGGAGCGGCGAAGTGATTTTCTGGATAACTCTCGTGGGGTATTCTCTCT 3894



Db 1343 -----AspGlySerHisThrLeuThrAlaThrValAsnAspLysAlaGlyAsnPro 1359  
 QY 3895 TTATCTACTGTGCCAACCTCATCA-----TCACCTCCCGCTGCTCCCAACAGTTAGTGAT 3948  
 Db 1360 AlAserThrThrHisAsnLeuAlaValAspLeuThrValProValLeuThrThrIleAsnThr 1379  
 QY 3949 GCTCGGAAGGGTCTATTTTCTGTAGAG-----ACTAGTTTGGAGATCTCAGGC 3999  
 Db 1380 IleAlaGlyAspAspIleIleAsnAlaThrGluHisGlyGlnAlaLeuValIleSerGly 1399  
 QY 4000 GTCAAAAAAGGG-----GTCATGTTCCGATAATAATCCCGGAATTC 4041  
 Db 1400 SerSerThrGlyGlyGluAlaGlyAspValValThrValThrLeuAsnSerLysThrTyr 1419  
 QY 4042 GGAACAGTTTTCGAGGTGAAGATAATAATATGCTGGTGGG-----GA 4086  
 Db 1420 ThrThrThrLeuAspAlaSerGlyAsnTrpSerValGly-ValProAlaAlaAspValTh 1439  
 QY 4087 GGCAGTGGGTCCGCTACACCATCAAGTAGCAGCTTTTACAGTTTAAAAAACTGTAAAGGGAA 4146  
 Db 1439 rAlaLeuGlySerGly-----ProGlnThrValThrAlaThrValThrAspAlaAlaGlyAs 1458  
 QY 4147 AGTTTCTTTCACAGATAAC 4165  
 Db 1458 n-----SerAspAsn 1461  
 RESULT 30  
 US-11-087-099-7521  
 ; Sequence 7521, Application US/11087099  
 ; Publication No. US20060041961A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abad, Mark S. et al.  
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement  
 ; FILE REFERENCE: 38-21(53450)B EP  
 ; CURRENT APPLICATION NUMBER: US/11/087,099  
 ; CURRENT FILING DATE: 2005-03-22  
 ; NUMBER OF SEQ ID NOS: 12464  
 ; SEQ ID NO 7521  
 ; LENGTH: 7465  
 ; TYPE: PRT  
 ; ORGANISM: *Magnetococcus* sp. MC-1  
 US-11-087-099-7521

Alignment Scores:  
 Pred. No.: 5.19e-09 Length: 7465  
 Score: 208.00 Matches: 313  
 Percent Similarity: 29.4% Conservative: 169  
 Best Local Similarity: 19.1% Mismatches: 617  
 Query Match: 2.6% Indels: 541  
 DB: 11 Gaps: 68

US-10-701-844-1 (1-4435) x US-11-087-099-7521 (1-7465)

QY 451 GGGGATATGCACAGCAATC-----ATGGTTCTCAAGAAATTTACGATGGG 498  
 Db 5670 GlyAlaAspThrAlaGluValSerIleAsnLeuThrAspGluThrLeuPheAlaGly 5689  
 QY 499 GAGACGTTAACTGATCATTTCCCTACTACTGTTATAGAGATCCGAGTGGGACTACTGTT 558  
 Db 5690 AspSerValAspIleThrPheSerArgAla----- 5699  
 QY 559 TTTTCTCGCAGAGATTAAACATTTAAATAATCTTCAATTTATGAGCTTTGCCCTTTA 618  
 Db 5700 ---LysThrGlyAsnLeuThr-----SerIleThrSerThrGlyLeu 5712  
 QY 619 AGTTGTTTGGGAATTTATTAGGAGTTTACTGTTTATGAGGAGACACACTCGTTCACT 678  
 Db 5713 TyrPheGlyGlyAsn-----AspAlaLeuValValGlyAspGly----AspValThr 5728  
 QY 679 TTC-----GAGACATACGGACTTCTCAAAATGGGGCA----- 711  
 Db 5729 PheIleGlyGlyLysAspAsnAspSerLeuIleValAspSerThrGlyAlaThrAla 5748

QY 711 ----- 711  
 Db 5749 GluValLeuAlaGlyThrAlaGlyAspProProAsnGlnArgPheAlaPheGlyAspAsn 5768  
 QY 712 -----GCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAG--- 753  
 Db 5769 ValTyrLeuGluTyrSerLeuAspAlaGlyLysGluAlaLeuIleThrIleAsnSer 5788  
 QY 754 -----GGTTTAAAGAAATATTCTCTTTTCCAAATTCGAATTC 789  
 Db 5789 GluAspThrThrAlaSerThrGlyAspAspSerLeuLeuIleGlyProThrGlySer 5808  
 QY 790 TTACTTGGCGTA-----CTCCCTGCTCAACGACTAAT 822  
 Db 5809 LeuAlaGlyValMetGlyGluValIleAlaValGlyGlyMetGlyAlaAspThrIleAsn 5828  
 QY 823 AAGGTAGCCAGACTCCGACGACACATCTACACGCTCTAATGGTACTATT--- 873  
 Db 5829 ValSerGlyLeuThrAlaAspThrValAlaAlaGlyAspAsnIleThrArgSer 5848  
 QY 874 -----TATTCTAAACACAGATCTTTTGTACTC-----AATAAT 906  
 Db 5849 ThrGluGlyTyrLeuLysThrMetArgSerLeuLeuProValGluGlyValAspAsn 5868  
 QY 907 GAGAGTTCTCTCATTTCTATAGTAATTTAGTC-----TCTGAGATGGGGAGCT 954  
 Db 5869 IleSerLeuSerAlaGlyAspHisValValIleGlyGlyMetSerGlyAspThrIleLeu 5888  
 QY 955 ATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTGTGCTTCCCAAGAAAT 1014  
 Db 5889 ThrGlySerGlyAsnThrIleAlaMetGlyAspSerGlySerLeuValPheAspSerAsp 5908  
 QY 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAA-----GTAGTCACACAGTTTC 1059  
 Db 5909 ThrAlaProValTyrTrpGlySerValAspSerThrAspAspProThrLeuThrGluAla 5928  
 QY 1060 TCTGCTATGGCT-----AACGAGGCTCTTATTGCTCCCTTTGTAGCGAATGTGCA 1107  
 Db 5929 GlnGluGlnAlaValLysArgValAsnArgTrpGlyIleLeuGlnThrLeuSerSerLeu 5948  
 QY 1108 GGAGTAAGAGGGGAGGATGCTGCTGTTTCAGGATGGGACGAGGAGGTGTCATCATCT 1167  
 Db 5949 GlyThrThrAsnGlyGlyHisAspThrMetThrLeuGlyGlnLysGlyThrSer--- 5966  
 QY 1168 ACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACTGCGGTAGAGTTTGATGGG 1227  
 Db 5967 -----PheAspGly 5969  
 QY 1228 AACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTCTGTAATAT 1287  
 Db 5970 AsnLysValAlaValGlyGly----- 5976  
 QY 1288 GGAACAACTGTTTCTCAACATGTTGCTTCTCTGTTTACATGCTGCTGAAGCAACCA 1347  
 Db 5977 -----IleGlyAsnAspThrIleThrLeuTyrGlyThrArgSerTyrVal 5991  
 QY 1348 ACAAGTGGACAGCTTCTAATACGATTAATAATACGAGATGGAGAGCTACTTCTGT 1407  
 Db 5992 GluThrAspAlaAlaGlyAsnAlaValGluAsnValAspGlySerAsnThrIleValVal 6011  
 QY 1408 AAGAATGGTGGCAAGCAGGATCCAATACTCTGGATCAGTTTCTTGTAGGAGGGA 1467  
 Db 6012 ThrGluArgAlaValAlaGlyAspAsn-----AlaThrMetSerPheGlyAlaAspGly 6029  
 QY 1468 GTAGTT-----TTCCTTAGTAGCAATGTAGCTCTGGGAAGGGGAGGCTATT 1515  
 Db 6030 LeuIleThrAspPheAlaThrGlyAspSerAsnIleAlaThrAlaGlyAsnAspThrIle 6049  
 QY 1516 TATGCCAAAAGCTCTCGTTGCTACTGTCGCTGTACAAATTTTAAAGATATGCT 1575  
 Db 6050 -----LysValAspValSerAsn---AspProValIleAspProAspTyrValAsp 6065  
 QY 1576 AATGATGGTGGAGCGATT-----TATTTA-----GGAGAA 1605

```
Db 6066 GlyAspIleGlyThrLeuAspMetTyrLeuThrThrGluPheAsnValLeuAlaGlyGly 6085
Qy 1606 TCTGGA-----GAGCTCAGTTTATCTGCTGATTATGGAGATATTATTTTCGAT 1653
Db 6086 LeuGlyAsnAspValLeuGluValGlySerThrGlyMetArgGluValIleSerGly 6105
Qy 1654 GGGAAATCTT-----AAAAGAACAGCCAAAGAGATGCTGCGAGTTTAAAGCGCTA 1704
Db 6106 AspAsnLeuLeuTyrGlnArgGlyThrAlaSerGlnAsnTyrAsnHisLeuLeuTyrAlaGlu 6125
Qy 1705 ACTGTCCTCACAAGCC-----ATTTGATGCGATCGGAGGAGAA--- 1746
Db 6126 ValIysSerSerSerGlyGlyThrAspThrIleIysThrGlyAsnGlyGluLeu 6145
Qy 1747 -----ATAACGACATTAAGAGCTAAAGACGCGGATCAG 1779
Db 6146 IlePheGlyGlyValGlyAsnAspThrLeuThrValLeuThrAlaValAspHisSer 6165
Qy 1780 ATTCTCTTTAATGATCCCATCGAGATGCGAAACGGAATAACAGCCAGCGAGCTTCC 1839
Db 6166 IleValPheGlyAspAlaGlyIysValAlaTyrAspThrAsnAlaThrGlyAlaValSer 6185
Qy 1840 AAATCTTCAAAAT-----AACGATGCTGAAGCA-----TACACAGG 1878
Db 6186 LeuIleAlaSerIleProGluThrAspAspGlyAsnAspThrValThrValIysAsnGly 6205
Qy 1879 GATATTGTTTTGCTAAAGAGCAGTACT-----TTGTACCAAAATGTTACGATAGAG 1932
Db 6206 AspLeuTyrLeuPheGlyGlyGlyThrAspAlaLeuThrValSerAlaAsnAspThr 6225
Qy 1933 CAAGGAAGATGTTCTTGCGAAAGGCAAAATATCAGTGAATTCCTAAGTCAGACA 1992
Db 6226 AlaThrArgValValMetGlyAspSerGlyGlnValAsnMetThrLeuTyrSerTyrThr 6245
Qy 1993 GGTGGAGTCTGTATATGGAGCTGGAGTACATGGGATTTGTAACTCCACCAACCA 2052
Db 6246 AspAla-----AspGlyValThr----- 6251
Qy 2053 CAACAGCCTCCTCGCGCTAAATCAGTTGATCAGCTTCCAAATCTGCATTTGCTCTTCT 2112
Db 6252 -----GlnThrIleGlyValProAsnPhe---IleGlnValAsn 6263
Qy 2113 TCTTTGTTAGCAACAATGCACTTACGAATCCTCTACCAATCTCCAGCGCAAGATCT 2172
Db 6264 AspAlaSerThrThrAsnGluSerThrArgIleAspThrPheThrLeuProSerIysGly 6283
Qy 2173 CATCTCGAGTCATTGGTAGCAACAATGCTGCTTCTGTACAAATAGTGGGCTATCTTT 2232
Db 6284 AspAsnLeuIleIleGly-----LeuGlyThrAspLeuLeuSerGly----- 6298
Qy 2233 TTTGAGGATTTGGATGATACAGCTTATGATAGTATGATTGGCTAGTTCTTAACAAA 2292
Db 6298 ----- 6298
Qy 2293 ATCAATGCTCGAAATTACAGTTAGGACTAAGCCC-----CCAGCTAATGCCCATCA 2346
Db 6299 -----AlaGlnLeuGlyValAspSerIleIleProGlyThrGlySerIle 6313
Qy 2347 GATTTGACTCTAGGATGAGATGCTTAAGTAGGC----- 2382
Db 6314 GlnValSerLeuAlaAsnSerAlaAspProThrGlyAlaTyrSerIysValValSerVal 6333
Qy 2383 -----TATCAAGGAAGCTGGAGCTGCGTGG 2409
Db 6334 ValValLeuGlyGlnTyrAsnGluMetGlyIleThrAspGlyThrTyrValLeuSerGln 6353
Qy 2410 GAT-----CCTAATACA-----GCAAAATAATGGCTTTATCTATCTGAAA 2448
Db 6354 AspGlyValAspThrArgProGluThrThrAlaThrGlySerGlyGluValThrGluAsp 6373
Qy 2449 GCTACATGACTAAATCTGGTAT-----AATCCTGGGCTGAGCGA 2490
```

```
6374 GlyThrGlnSerValThrGlyTyrLeuGlyLeuAspAlaLeuAsnGlyGlyLeuAlaIle 6393
2491 GTAGCTTCTTTGGTCCAAATAGTTTATGGGATCCATT----- 2529
6394 PheThrGluAlaThrThrThrLeuLeuTyrGlyThrLeuThrValIysSerAspGlyGly 6413
2530 -----TTAGATATACGATCTGCGCATTCAGCAATTCAGCA----- 2565
6414 TrpThrTyrThrLeuAsnAsnSerSerAlaValGlnAlaLeuMetSerGlyAsnAsp 6433
2566 -----AGTGTGATGGGCGCTCTTATTTGTCGAGGATTTATGG 2601
6434 ArgProGluThrPheGlnValAsnThrThrAspGlyGlnGlnThrThrValThrIleArg 6453
2602 GTTCTGAGAGTTTCCAAATTTCTTCTATCATGACCGCATGCTTTAGGTCAAGGATATCGG 2661
6454 ValIysGlyAlaAspAsp-----ValSer 6461
2662 TATATTAGTGGGTTTATCTCTTAGGACCAACTCTTCTTGGATCATGATGTTTGGT 2721
6462 ThrLeuThrGlySerSerSerAlaSerLeuAsnGluAspAspThrSerAlaValSerGly 6481
2722 CTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATATAGTAGTGTGCTTCCCAATCAT 2781
6482 ---ThrLeuThrValValAspAlaAspValIleAspAlaThrValThrAlaAlaThrSer 6500
2782 CATGCTTCATPAGGATCGTTTATCTATCTATCCCAA----- 2817
6501 ValGlyThrTyrGlyThrPheSerValGlySerAsnGlyValTrpSerTyrGlnLeuAsp 6520
2818 -----CAAGCTTATGTGGATCCCTATTTGTCGAGATGCGTTATC 2859
6521 AsnSerIysAlaValValGlnGlyLeuThrGlnGlyGlnValSerGlySerPheThr 6540
2860 CGTGTACTACGGTTTGGGAATCAGCATATGAAACCTCATATACATATTCGACAGGAG 2919
6541 ValHisThrSerAspGlyGlySerGlnSerLeuThrMetAsnIleThrGlyArgGlnAsp 6560
2920 AGCGATGTTCTGGGATTAATPACTGTCTGCTGGAGAGATTGGAGCGGATTAACCGATT 2979
6561 SerAlaVal-----ValGlySerGlyThrGlySer 6570
2980 GTGATTACTCCTAAGCTCTATTTGAATGAGTTGCTCTTCTGTCAGACTCAGTTT 3039
6571 ValThrGluAspSerThrLeuSerSerSerGlyMetLeuSerValSerAspAlaAspSer 6590
3040 TCTTATGCGGATCATGAATCTTTACAGCAAGAGC-----GAT 3078
6591 GlyGluAlaAlaTyrValAlaSerThrGlnThrGlySerTyrGlySerPheSerLeuAsp 6610
3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTTAATCTATCAAGTTCTCTGTTGGAGTGAAG 3138
6611 SerSerGlyAlaTyr-----SerTyrGlnLeuAsn----- 6620
3139 TTTGATCATGTTCTTAGTACATCTCTAATAATATAGATCTTTATGCGGCTTATATCTGT 3198
6621 -----AsnSerAlaSerGlnValGlnAlaLeuIleMetGlyGlnGlnValSer 6636
3199 GATGCTTAT-----CGCACCATCTCTGCTACTGAGACA-----ACGCTCTCTTA 3240
6637 GluSerPheThrValAlaThrValAspGlyThrAlaSerSerValValValThrValVal 6656
3241 TCCCATCAAGACATGAGCAACAGATCCCTTTTCAATTTAGCA-----AGACATGGA 3291
6657 GlyAlaGln-----AspAlaAlaGlnLeuGlyGlyValGlnSerGlyGly 6671
3292 GTTGTGGTTAGGGA-----TCTATGATGCTCTCTTAACCAAGTAATATAGATATAT 3345
6672 ValValGluAspGlyGlnLeuSerThrAlaGlyGlnLeuThrIleSerAspAlaAspAla 6691
3346 GCCCATGGAAGATGAGTATCGAGATGCTTCCGAGGCTATGTTTGTAGTCAGGAACT 3405
6692 GlyGlnSerLeuTyrGlnValAlaSerThrAspGlyMetTyrGlyHisPheAlaMetAsp 6711
```



QY 664 ---GGACACTCGTTGACCTTCGAGACATACCGACTTCTACA----- 702  
Db 120 AspSerGlnProValThrPheValAlaAspLysThrSerAlaLeuValValLeuGlnIle 139  
QY 703 -----AATGGG-----GCAGCTCTAAGTAATAGCGCTGCT 732  
Db 140 SerLysAsnGluIleThrGlyAsnGlyValAspSerAlaThrLeuThrAlaThrValLys 159  
QY 733 GATGACTGTTTACTATTGAGGGTTTTAAAGAAATATCTCTTTCCAATTGCAATTCAATTA 792  
Db 160 AspGlnPheAspAsnGluValAsnAsnLeuProValThrPheSerThrAlaSerSerGly 179  
QY 793 CTTCGGCTACTGCTGCT---GCAACGACTAATAAG----- 825  
Db 180 LeuThrLeuThrProGlyGluSerAsnThrAsnGluSerGlyIleAlaGlnAlaThrLeu 199  
QY 826 -----GGTAGCCAGACTCCGACGACACATCTACACCGCTCTAATGGTACT 870  
Db 200 AlaGlyValAlaPheGlyGluGlnThr---ValThrAlaSerLeuAlaAsnAsnGlyAla 218  
QY 871 ATTTATCTTAAACAGATCTTTTGTGTTACTCAATAATAGAAAGTTCTCATTTCTATAGTAAT 930  
Db 219 SerAspAsnLysThrValHisPheIle----- 227  
QY 931 TTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAAACGGTTCAAGGAATTAGC 990  
Db 228 -----GlyAspThrAlaAlaAlaLysIleIleGluLeuThrProValProAspSer 244  
QY 991 AAGCTTTGTCTCTCAAGMAAACTGCTCAAGCTGATGGGGAGCTGTGCAAGTAGTC 1050  
Db 245 IleIleAla-----GlyThrProGlnAsnSerSerGlySer-----ValIle 258  
QY 1051 ACCAGTTTCTGCTATGCTTAACAGAGCTCTATTGCTTTGTA----- 1095  
Db 259 ThrAlaThrValValAspAsnAsnGlyPheProValLysGlyValThrValAsnPheThr 278  
QY 1096 -----GCGAATGTTGCGAGGATAGAGGGGAGGATGCTGCTGTTTCAGGATGGGCAG 1149  
Db 279 SerAsnAlaAlaThrAlaGluMetThrAsnGlyGlyGlnAlaValThrAsnGluGlnGly 298  
QY 1150 CAGGAGGTGTCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACT 1209  
Db 299 LysAlaThrValThrTyThrAsnThrArgSerSerIleGluSerGlyAlaArgProAsp 318  
QY 1210 CGGTPAGATTTGATGGGAACGTACCCGAGTAGGAGGAGGATTTACTCTACGGGAAC 1269  
Db 319 ThrValGlu----- 321  
QY 1270 GTTGCTTTCTGTAATAATGGAAAA-----ACCTGTTTCTCAACAATGTT 1314  
Db 322 ---AlaSerLeuGluAsnGlySerSerThrLeuSerThrSerIleAsnValAsnAlaAsp 340  
QY 1315 GCTTCTCTGTTTACATGCTGCTAAGCAACCA-----ACAGTGGACAG 1359  
Db 341 AlaSerThrAlaHisIleuThrLeuLeuGlnAlaLeuPheAspThrValSerAlaGlyAsp 360  
QY 1360 GCTTCTAAT-----ACGAGTAATAATTACGAGATGGAGGA----- 1395  
Db 361 ThrThrAsnLeuTyrIleGluValLysAspAsnTyrglyAsnGlyValProGlnGlnGlu 380  
QY 1396 -----GCTATCTTCTGTAAGATGTTGGCGCAGCGGATCCAAATCACTCTGATCA 1446  
Db 381 ValThrLeuSerValSerProSerGluGlyValThrProSerAsnAsnAlaIleTyThr 400  
QY 1447 GTTTCCTTTGAGGAGGGAGTAGTTTCTTTTAGCAATGTTAGCTGCTGGGAAAGGG 1506  
Db 401 ThrAsnHisAspGlyAsn-----PheTyralaSerPheThrAlaThrLysAla 416  
QY 1507 GGAGCTATTATGCCAAAAGCTCTGGTGTCTAACTGTGGCCCTGTATCAA----- 1557  
Db 417 Gly---ValTy-GlnValThrAlaThrLeuGluAsnGlyAspSerMetGlnGlnThrVal 435  
QY 1558 ---TTTTTAAGGAATATCGCTAATGATGTTGGGCGGATTTATTTAGGAGAAATCTGGAGAG 1614

Db 436 ThrTyrrValProAsnValAlaAsn----- 445  
QY 1615 CTCAGTTTATCTGCTGATATGGAGATATTATTTCATGGGAATCTTTAAAGAACAGCC 1674  
Db 446 IleSerLeuAlaAlaSerLysAspProValIleAlaAsnAsnAsnLeuThrThrLeu 465  
QY 1675 AAAGAGATGCTCGCGATGTTTAATGGGTAACTGTGCTCTCACAGCCATT----- 1725  
Db 466 ThrAlaThrValAlaAspThrGluGlyAsnAlaIleAlaAsnSerGluValThrPheThr 485  
QY 1726 -----TCGATGGATCGGAGCGGAAAAATAACGACATTA 1758  
Db 486 LeuProGluAspValArgAlaAsnPheThrLeuGlyAspGlyGlyValValThr--- 504  
QY 1759 AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818  
Db 504 ----- 504  
QY 1819 AACGAGCCAGCGAGCTCTTCCAAACTTCTAAAAATTAAACGATGGTGAAGGATACACAGG 1878  
Db 505 AspThrGluGlyLysAlaLysValThrLeuLysGlyThrLysAlaGlyAlaHisThrVal 524  
QY 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGGA 1938  
Db 525 ThrAlaSerMetAlaGlyLysSer----- 533  
QY 1939 AGGATTGTTCTCTGTAAGAGCAAAATTAATCAGTGAATCTCTAAGTCAGACAGGTGG 1998  
Db 534 -----GluGlnLeuValValAsnPheIleAlaAspThr----- 544  
QY 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCACCAACAG 2058  
Db 544 ----- 544  
QY 2059 CCTCTGCGCTAATCAGTTGATCAGCTTTCCTCAATCTGCATTTGTCTCTTCTCTTTG 2118  
Db 545 -----LeuThrAlaGlnValAsnLeuAsnValThrGluAspAsnPhe 558  
QY 2119 TTAGCAACAAT-----GCAGTTACGAATCTCTCTACCAAT 2154  
Db 559 IleAlaAsnAsnValGlyMetThrArgLeuGlnAlaThrValThrAspGlyAsnGlyAsn 578  
QY 2155 CTTCAGCGCAAGATTCTCATCTGTCAGTCACTGGTAGCACAACTGCTGTTCTGTATCA 2214  
Db 579 ProLeuAlaAsnGlu-----AlaValThr 586  
QY 2215 ATTATGGGCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGTATGATTGG 2274  
Db 587 PheThr----- 588  
QY 2275 CTAGGTTCTAATCAAAAAATCAATGTCCTGAAATTTACAGTTAGGAGCTAAGCCCCAGCT 2334  
Db 589 -----LeuProAla 591  
QY 2335 AATGCCCATCAGATTGACTCTAGGAAATAGATGCCCTAAGTATGGCTATCAAGGAAGC 2394  
Db 592 AspValSerAlaSerPheThrLeuGly-----GlnGlyGly 603  
QY 2395 TGGAGCTTGGCTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACA 2454  
Db 604 SerAlaIle-----ThrAspIleAsnGlyLysAlaGluValThrLeuSer 618  
QY 2455 TGGACTAAACCTGGG-----TATAATCTCGGGCT 2484  
Db 619 GlyThrLysSerGlyThrTyrrProValThrValSerValAsnAsnTyrrGlyValSerAsp 638  
QY 2485 GAGCAGTAGCTCTTGTGTTCCAAATAGTTTATGGGATCCATTTTATAGATATAGATCT 2544  
Db 639 ThrLysGlnValThrLeuIleAlaAspAlaGlyThrAlaLysLeuAlaSerLeuThrSer 658  
QY 2545 GCGCATTCAGCAATTCAGCAAGTGTGATGGGCTCTTATTGTCGAGGATTATGGGTT 2604

Db 659 ValTyrSerPheValValSerThrThrGluGlyAlaThrMetThrAlaSerVal----- 676  
Qy 2605 TCTGGAGTTTCGAATTTCTTATCATGACCGCGAGTCTTAGGT-----CAGGA 2655  
Db 677 -----ThrAspAlaAsnGlyAsnProValGluGly 686  
Qy 2656 TATCGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCC----- 2697  
Db 687 IleLysValAsnPheArgGlyThrSerValThrLeuSerSerThrSerValGluThrAsp 706  
Qy 2698 -----TACTTTGGATCATCGATGTTTGGTCTAGCAATTT----- 2730  
Db 707 AspArgGlyPheAlaGluIleLeuValThrSerThrGluValGlyLeuLysThrValSer 726  
Qy 2731 -----ACCGAAGTATTTGGTAGTACTAAGATTATGTAGTGT 2769  
Db 727 AlaSerLeuAlaAspLysProThrGluValIleSerArg----- 739  
Qy 2770 CGTTCCAATCATCATGCTGGATAGGATCGTTTATCTATCTATACCAACAAAGCTTTATGT 2829  
Db 740 LeuLeuAsnAlaLysAlaAspIleAsnSerAlaThrIleThrSerLeuGluIleProGlu 759  
Qy 2830 GATCTCTATTCTTCGGA---GATGCGTTTATCCGTGCTAGTACGGG-----TTTGG 2880  
Db 760 GlyGlnValMetValAlaGlnAspValAlaValLysAlaHisValAsnAspGlnPheGly 779  
Qy 2881 AATCAGCATATGAAACCTCATATACATTTCGACGAGGAG-----AGCATGTTCTGTTGG 2934  
Db 780 AsnProIleLeuAsnGluSerValThrPheSerAlaGluProProGluHisMetThrIle 799  
Qy 2935 GATAATACTGCTGCTGGAGAGATGGCGGGATTACCGATTGTGATTATCCATCT 2994  
Db 800 SerGlnAsnIleValSerThrAspThrHisGlyIleAlaGluValThrMetThrPro--- 818  
Qy 2995 AAGCTCTATTGTAATGAGTTGCGCTTCGTCGAAGCTGAGTTTCTTATGCGCATCAT 3054  
Db 819 -----GluArgAsnGlySerTyrMetValLysAlaSerLeuAlaAsnGly 833  
Qy 3055 GAATCTTTTACAGAGAA-----GGCGATCAAGCTCGGGCATTCAGAGCGGA 3102  
Db 834 SerSerTyrGluLysAspLeuValValIleAspGlnLysLeuThrLeuSerAlaSerSer 853  
Qy 3103 CATCTCTAATCATCATGTTCTCTGTTGGAGTGAAGTTTCAT-----CGATGTTCT 3153  
Db 854 ProLeuIleGlyValAsnSerProThrGlyAlaThrLeuThrAlaThrLeuThrSerAla 873  
Qy 3154 AGTACACATCTTAATAATAGTCTTATGCGCGCTTATATCTGTCATGCTTATCGCACC 3213  
Db 874 AsnGlyThrProValGluGlyGlnValIleAsnPheSerValThrProGluGlyAlaThr 893  
Qy 3214 ATCTCTGGTACTGAGACACGCTCTATCCATCAAGAG----- 3252  
Db 894 LeuSerGlyGlyLysValAlaGThrAsnSerSerGlyGlnAlaProValValLeuThrSer 913  
Qy 3253 -----ACATGGACA---ACAGATGCCCTTTCATTAGCAACATGAGTTGTG 3297  
Db 914 AsnLysValGlyThrTyrThrValThrAlaSerPheHis-----AsnGlyValThr 930  
Qy 3298 GTTAGAGGATCATGTATGCTTCTCAACAAGTAATATAGAATATATAGCATATGCGCATGAAGA 3357  
Db 931 IleGlnThrGlnThrIleValLysValThrGlyAsnSerSerThr---AlaHisValAla 949  
Qy 3358 TATGAGTATCGAGATGCTTCTCGAGGCTATGGT-----TTGAGT 3396  
Db 950 SerPheIleAlaAspProSerThrIleAlaAlaThrAsnSerAspLeuSerThrLeuLys 969  
Qy 3397 GCAGGAAGTAGAGTCGGTTCTTAAATATGTTAGTATAGTATTAGTGTAGCATGCCCT 3456  
Db 970 Ala-ThrValGluAspGlySerGlyAsnLeu---IleGluGly-----LeuThrValTyr 986  
Qy 3457 TTTTCTTTGAGATCATCATATTGTTTTTTTGTGTTGTTGTTCTATTCG----- 3510  
Db 987 PheAlaLeuLysSerGlySer-----AlaThrLeuThrSerLeuThrAlaVal 1002

Qy 3511 -----TATGATTTCGAGCTCTCT-----CAAGTGTTAACGCCCTPAATGTAAACC 3555  
Db 1003 ThrAspGlnAsnGlyIleAlaThrThrSerValArgGlyAlaIleThrGlySerValThr 1022  
Qy 3556 ACTCCTTTAAGGAGACGATGTTTACTTGAATGGAGACTCGGCTTTTGTCAATGTC--- 3612  
Db 1023 Val-----SerAlaValThrThrAlaGlyMetGlnThrValAspIleThr 1038  
Qy 3613 ---TATCGACGAGCTGAAGAAGGT-----TCGATTATCTCAGCTAAATGGCGACAATTTA 3663  
Db 1039 LeuValAlaGlyProAlaAspAlaSerGlnSerValLeuLysAsnAsnArgSerSerLeu 1058  
Qy 3664 -----ACGATTACCGGACAAACCATCATATTATCATTTACAGATTCTCAAGGG 3711  
Db 1059 LysGlyAspPheThrAspSerAlaGluLeuHisLeuValLeuHisAspIleSerGlyAsn 1078  
Qy 3712 CCA-----GTTCTTCAAAATATCCCTTCATTTCAGCAGGAGACACTTACTCTGACA 3765  
Db 1079 ProIleLysValSerGluGlyLeuGluPheValGlnSerGlyThrAsnAlaProTyrVal 1098  
Qy 3766 GATTTTTCGAGTCTGATGTTCTCGAAAATGTTCTTCGGAGAGAAAAGGAATGATCTCC 3825  
Db 1099 GlnValSerAlaIleAspTyrSerLysAsnPheSerGlyGluTyrLysAlaThrValThr 1118  
Qy 3826 GGGAAAAACCGTGAGTATTTTCGGAGACGCGAAGTATTTCTGGATTAACCTCCGTGGGG 3885  
Db 1119 GlyGlyGlyGluGlyIleAlaThrLeuIleProValLeu-----AsnGlyValHis 1135  
Qy 3886 TATTCCTCTTATCTACTGTGCAACCTCATCATCACTCCGCTGCTCCACAGTTAGT 3945  
Db 1136 GlnAlaGlyLeuSerThr-----ThrIleGln 1144  
Qy 3946 GATGCTCGAAAGGCTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAA 4005  
Db 1145 PheThrArgAlaGluAspLysIleMetSerGlyThrValLeu---ValAsnGlyAlaAsn 1163  
Qy 4006 AAAGGGTCACTTCGATAATAATGCCGGAATTTTCGAAACAGTATTTTCGAGGTAAGAAT 4065  
Db 1164 LeuProThrThrThrPheProSerGlnGlyPheThrGlyAlaTyrTyrGlnLeuAsnAsn 1183  
Qy 4066 AATAATAATGCT---GGTGGTGGAGCAGTGGGTTCCGCTACACATCAAGTACGACTTT 4122  
Db 1184 AspAsnPheAlaProGlyLysThrAlaAlaAspTyrGluPheSer-SerSerAlaSerTr 1203  
Qy 4123 TACAGTTAAAACTGTAAAGGGAAGTCTTTTTCACAGATAACGTAGCCTCT 4174  
Db 1203 pValAspValAspAlaThrGlyLysValThrPhe---LysAsnValGlySer 1219

## RESULT 32

US-11-052-554A-280  
; Sequence 280, Application US/11052554A  
; Publication No. US2005028886A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 280  
; LENGTH: 1588  
; TYPE: PRN  
; ORGANISM: Escherichia coli 0157:H7  
US-11-052-554A-280  
Alignment Scores:

```

Pred. No.: 4,69e-09 Length: 1588
Score: 205.00 Matches: 293
Percent Similarity: 30.9% Conservative: 203
Best Local Similarity: 18.3% Mismatches: 577
Query Match: 2.6% Indels: 531
DB: 11 Gaps: 69

US-10-701-844-1 (1-4435) x US-11-052-554A-280 (1-1588)
QY 523 TATACCTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTCTCGCAGGAGATTAAACATTA 582
Db 16 TyrThrValThrSerGluThrAlaLysSerArgGlyLysLysSerGlyArgSerLysLeu 35
QY 583 AAAAATCTTGCAATCTATTGCGACCTTTCCTTTAAAGTTGTTTGGGAACCTATTAGG 642
Db 36 -----LeuIleSerAlaLeuValAla -----GlyGlyMetLeuSer 47
QY 643 AGTTTTTACTGTTTATAGGAGAGGACACTCGTTTGACTTTTCGAGAACATACGGACTTCTACA 702
Db 48 SerPheGlyAlaLeuAlaAsnAla -----GlyAsnAsp 58
QY 703 AATGGGCGAGCTCTA-----AGTAATAGCGCTGCTGATGGACTGTTTACTATT--- 750
Db 59 AsnGlyGlnGlyValAspTyrGlySerGlySerAlaGlyAspGlyTyrValAlaIleGly 78
QY 751 GAGGTTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTTACTTGGCGTACTGCTGCT 810
Db 79 LysGlyAlaLysAlaAsnThrPheMetAsn----- 88
QY 811 GCAACGACTAATAAGGGTAGCGACACTCCGACGACACATCTACACCGCTTAATGGT--- 867
Db 89 -----ThrSerGlySerSerThrAlaValGlyTyrAspAlaIleAlaGluGlyGln 105
QY 868 -----ACTATTTATTTAAAAAGATCTTTTGTACTCAATAATGAGAAGTTCTCA 918
Db 106 TyrSerSerAlaIleGlySerLysHis-----AlaIleGlyLysAlaSerMetAla 123
QY 919 TTCTATAGTAATTTAGTCTCT---GGAGATGGGGA---GCTATAGATGCTTAAGAGCTTA 972
Db 124 PheGlyValSerAlaIleSerGluGlyAspArgSerIleAlaLeuGlyAlaSerSerTyr 143
QY 973 ACGGTT-----CAAGGAATTAGCAAGCTTTGT 999
Db 144 SerLeuGlyGlnTyrSerMetAlaLeuGlyArgTyrSerLysAlaLeuGlyLysLeuSer 163
QY 1000 GTCTTCCAGAAAATACTCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAAGTTTC 1059
Db 164 IleAlaMetGlyAspSerSerLysAlaGlu---GlyAlaAsnAlaIleAlaLeuGlyAsn 182
QY 1060 TCTGTATGGCTAACGAGCTCTATTGCTTTTGTAGGAAATGTTGCGAGGAGTAAGA--- 1116
Db 183 AlaThrLysAlaThrGluIleMetSerIleAlaLeuGlyAspThrAlaAsnAlaSerLys 202
QY 1117 -----GGGGGGGATGCTGCTGTTTCAGGATGGCGGAGGAGTG 1158
Db 203 AlaTyrSerMetAlaLeuGlyAlaSerSerValAlaIleGlyAsnAsnAlaIleAlaIle 222
QY 1159 TCATCATCTCTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTCGGTAGAG 1218
Db 223 GlyAlaGluThrGluAlaAlaGluAsnAlaThrAlaIleGlyAsnAsnAlaLysAlaLys 242
QY 1219 TTTGATGGGAACGTAGCCGAGTAGAGAGGAGTTTACTCTACGGGAACCTGCTTTC 1278
Db 243 GlyThrAsnSerMetAlaMetGlyPheGlySerLeuAlaAspLysValAsnThrIleAla 262
QY 1279 CTGAATAATGGAAAACCTGTTTCTCAACAATGTTGCTTCTCCTGTTTACATGCTGCT 1338
Db 263 LeuGlyAsnGlySerGlnAlaLeuAlaAspAsnAla-----IleAlaIle 277
QY 1339 AAGCAACCAACCAAGTGGACAGCTCTTAATACGAGTAATAATTAACGAGATGGA---GGA 1395
Db 278 GlyGln-----GlyAsnLysAlaAspGlyValAsp 287

```

```

QY 1396 GCTATCTTCTGTAAGAATGGTGGCAGCA----- 1425
Db 288 AlaIleAlaLeuGlyAsnGlySerGlnSerArgGlyLeuAsnThrIleAlaLeuGlyThr 307
QY 1426 -----GGATCCAAATAACTCT----- 1440
Db 308 AlaSerAsnAlaThrGlyAspLysSerLeuAlaLeuGlySerAsnSerSerAlaAsnGly 327
QY 1441 ---GGATCAGTTCTCTTTGATGGAGAGGAGTAGTTCTTTCTTAGTAGCAATGTAGCTGCT 1497
Db 328 IleAsnSerValAlaLeuGlyAlaAspSerIleAlaAspLeuAspAsnThrValSerVal 347
QY 1498 GGG-----aaGGGGGAGCTATT----- 1515
Db 348 GlyAsnSerSerLeuLysArgLysIleValAsnValLysAsnGlyAlaIleGlySerAsp 367
QY 1516 -----TATGCCAAAAAGCTCTCGTTGCTTAAC--- 1542
Db 368 SerTyrAspAlaIleAsnGlySerGlnLeuTyrAlaIleSerAspSerValAlaLysArg 387
QY 1543 -----TGTGGCCCTGTACAAATT 1560
Db 388 LeuGlyGlyAlaAlaValAspValAspAspGlyThrValThrAlaProThrTyrAsn 407
QY 1561 TTAAGGATATCGCT-----AATGATGGTCGAGCGATTATTATTAGGAGATCTGGAGAG 1614
Db 408 LeuLysAsnGlySerLysAsnAsnValGlyAlaAlaLeuAlaValLeuAspGluAsnThr 427
QY 1615 CTC----- 1617
Db 428 LeuGlnTyrAspGlnThrLysGlyLysTyrSerAlaAlaHisGlyThrSerSerProThr 447
QY 1618 -----AGTTTATCTGCTGATTTATGGAGATATTATT 1647
Db 448 AlaSerValIleThrAspValAlaAspGlyThrIleSerAlaSerSerLysAspAlaVal 467
QY 1648 TTCGATGGGAATCTTAAAGAACACAGCAAGATGCTGCGATGTTTAATGGCGTAAT 1707
Db 468 AsnGlySerGlnLeuLysAlaThrAsnAspAspValGluAlaAsnThrAlaAsnIleAla 487
QY 1708 GTGTCCTCACAAGCATTTTCGATGGATCGCGAGG-----AAAAATA 1749
Db 488 ThrAsnThrSerAsnIleAlaThrAsnThrAlaAsnIleAlaThrAsnThrThrAsnIle 507
QY 1750 ACGCATTTAGAGCTTAAAGCAGG-----CATCAGATTCTCTTTAATGATCCC 1797
Db 508 ThrAsnLeuThrAspSerValGlyAspLeuGlnAlaAspAlaLeuLeuTyrAsnGluThr 527
QY 1798 ATCGAGATGGCAACGGAAATAACCGCCAGCGAGTCTCCAACTTCTAAAATTAAC 1857
Db 528 LysLysAlaPheSerAlaAlaHisGlyGlnAspThrThrSerLysIleThrAsnValLys 547
QY 1858 GATGGTGAA-----GGATACACA 1875
Db 548 AspAlaAspLeuThrAlaAspSerThrAspAlaValAsnGlySerGlnLeuLysThrThr 567
QY 1876 GGGGATATTGT-----TTGCTAATGGAAGCAGTACTTTGTACCAA 1917
Db 568 AsnAspAlaValAlaThrAsnThrThrAsnIleAlaAsnAsnThrSerAsnIleAlaThr 587
QY 1918 AATGTTACGATAGACCAAGGAAGGATTCTTCTCGTGAAGGCAAAATATCATCGTGAAT 1977
Db 588 AsnThrThr-----AsnIleSer 593
QY 1978 TCTCTAAGTCAGACAGGTGGGAGTCTGTATATGGAAGCT----- 2016
Db 594 AsnLeuThrGluThrValThrAsnLeuGlyGluAspAlaLeuLysTyrAspLysAspAsn 613
QY 2016 ----- 2016
Db 614 GlyValPheThrAlaAlaHisGlyThrGluThrThrSerLysIleThrAsnValLysAsp 633
QY 2017 -----GGGAGTACATGGGATTTGTAACTCCACACACCACCAACAGCT 2061

```





QY 3874 AACTCCGCGGGTATTCTTCCTTATCTACTGTGCCAACCTCATCATCACTCCGCTGCT 3933  
 Db 1249 AenSerIleAlaIleGlyAenGlySerThrThrThrArgGlyAlaGlnThrAsnThr 1268  
 QY 3934 CCAACAGTTAGTGTCTCGAAGGGTCTATT-----TTTTCTGTGA-----CAG 3978  
 Db 1269 AlaTyAsnMetAspAlaProGlnAsnSerValGlyGluPheSerValGlySerAlaAsp 1288  
 QY 3979 ACTAGTTTGGAGATCTCAGGGCTCAAAAAGGGGTCTGTTCTCGATAATAATGCCGGGAAT 4038  
 Db 1289 GlyGlnArgGlnIleThrAsnValAlaAlaGly---SerAlaAspThrAspAlaValAsn 1307  
 QY 4039 TTCGGAACAGTTTTCGAGGTAAAGTAATAATAATGCTGGTGGGAGGAGTGGGTTTC 4098  
 Db 1308 ValGly-GlnLeu-----LysValThrAs 1315  
 QY 4099 CGCTACACCACTCAAGTACGACTTTTACAGTTTAAACCTGTAAGGAAAGTTTCTTTCAC 4158  
 Db 1315 pAlaGlnValSerGlnAsnThrGlnSerIleThrAsnLeuAspAsnArgValThrAsnLe 1335  
 QY 4159 AGAT-----AACGTAGCTCTTTCGCGAGCGGAGTGGTTTATAAGGCAATTGT 4206  
 Db 1335 uAspSerArgValThrAsnIleGluAsnGlyIleGlyAspIleValThrThrGlySerTh 1355  
 QY 4207 G---CTTTTCAAGACATGAAGGAGGACATATTTCTCGAGGGAACACAGCATACGATGA 4263  
 Db 1355 rLysTyThrPheLysThrAsnThrAspGlyVal----- 1365  
 QY 4264 TTTAAGGATCTTCTGCTACTTAATCAGATCAGATCAGATCAGATCAGGAGCGGGTGGAGG 4323  
 Db 1366 -----AspAlaSerAlaGlnGlyLysAspSer---ValAlaIleGlySerGI 1380  
 QY 4324 AGTTATTGCTCTCCAGATGATCTGTAAAGTTTGAAGCAATAAAGTTCTTATTGTTTT 4383  
 Db 1380 ySerIleAlaAlaAlaAspAsnSerValAlaLeu-----GlyThrGlySerValAlaTh 1398  
 QY 4384 TGATTACAC 4393  
 Db 1398 rGluGluAsn 1401

RESULT 33  
 US-11-052-554A-92  
 ; Sequence 92, Application US/11052554A  
 ; Publication No. US2005028866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sachdeva, et al.  
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
 ; FILE REFERENCE: 30853/40359A  
 ; CURRENT APPLICATION NUMBER: US/11/052,554A  
 ; PRIOR FILING DATE: 2005-02-07  
 ; PRIOR APPLICATION NUMBER: US 60/589,227  
 ; PRIOR FILING DATE: 2004-07-20  
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
 ; NUMBER OF SEQ ID NOS: 763  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 92  
 ; LENGTH: 2399  
 ; TYPE: PRT  
 ; ORGANISM: Helicobacter pylori J99  
 US-11-052-554A-92

Alignment Scores:  
 Pred. No.: 8,46e-09 Length: 2399  
 Score: 203.00 Matches: 218  
 Percent Similarity: 29.9% Conservative: 139  
 Best Local Similarity: 18.3% Mismatches: 340  
 Query Match: 2.6% Indels: 496  
 DB: 11 Gaps: 52

US-10-701-844-1 (1-4435) x US-11-052-554A-92 (1-2399)

QY 469 ATCATGTTCTCAAGGAATTTACGAT-----GGGGAGCGTTAACTGTATCA 516  
 Db 1629 LeuLeuSerProArgGlyLeuHieAspPheTrpGlnLysGlyTyPheAsnPhLeuSer 1648  
 QY 517 TTTCCCTATATCTGTTATAGGAGATCCGAGT-----GGGACTACTGTT 558  
 Db 1649 AsnGlyTyThrValPheValAsnAsnSerPheSerAsnAlaThrGlyGlySerLeuAsn 1668  
 QY 559 TTTTCTCGAGAGAGTTAAACATTAATAAATCTTCAAAATCTTATTTGCGAGCTTTGCTTTA 618  
 Db 1669 PheValAlaAsnLysSerIlePheAsnGlyAspAsnThrIleAsp----- 1684  
 QY 619 AGTTGTTTGGAACTTATAGGAGGTTTACTGTTTATAGGAGGAGGACACTCGTTGACT 678  
 Db 1685 -----PheSerLysTyThrGlnGlyAlaLeuIlePheAlaSerAsnGlyValSerAsnIle 1702  
 QY 679 TTCGAGAACATACCGACT-----TCTACAATGGGCGAGCTCTAAGTAATAGCGCTGCT 732  
 Db 1703 -----AsnIleThrThrLeuAsnAlaThrAsnGlyLeuSerLeu---AsnAlaGlyLeu 1719  
 QY 733 GATGAGCTGTTTACTATTAGGGTTTTAAAGAAATATCTCTTTTCCAAATTCGAAATTCATTA 792  
 Db 1720 AsnAsnValSerValGlnLysGlyGluIleCysIleAsnLeuAlaAsnCys----- 1736  
 QY 793 CTTCGCGTACTGCTGCTGCGAACGACTAATAAGGGTAGCCAGACTCCGACGACA---ACA 849  
 Db 1737 -----ProThrThrLysAsn 1741  
 QY 850 TCTACACGCTCTAATGGTACTATTATTCTAAAACAGATCTTTGTTACTCAATTAATGAG 909  
 Db 1742 SerSerProAlaAsnSerSerValThrProThrAsnGluSerLeuSerValHisAlaAsn 1761  
 QY 910 AAGTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGAGCTATAGATGCTTAAGAGC 969  
 Db 1762 AsnPheThrPheLeuGlyThrIleIleSer-----AsnGlyAlaIleAspLeuSerGln 1779  
 QY 970 TTAACGGTTCAGGAATTTAGCAAGCTTTGTGTCTTCCAAAGAAAT---ACTGCTCAAGCT 1026  
 Db 1780 ValThrAsnAsnSerValIleGlyThrLeuAsnLeuAsnGluAsnAlaThrLeuGlnAla 1799  
 QY 1027 GATGGGGAGCTTGTCAAGTAGTACACAGTCTTCTGCTATGGCTTAAAGAGGCTCTCTATT 1086  
 Db 1800 AsnAsn-----LeuThrIleThrAsnAlaPheAsnAsn----- 1810  
 QY 1087 GCCTTTGTAGCGAATGTTGCAGGAGTAAGAGGGGAGGATTTGCTGCTTCAGGATGG 1146  
 Db 1810 ----- 1810  
 QY 1147 CAGCAGGAGTGTCTATCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAAT 1206  
 Db 1811 -----AlaSerAsnSerThrAlaAsnIleAspGlyAsnPheThrLeuAsnGlnGln 1827  
 QY 1207 ACTGCGGTAGAGTTTGTATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTCAAGGG 1266  
 Db 1828 AlaThrLeuSerThrAsnAlaSerGlyLeuAsnValMetGlyAsnPheAsnSerTyGly 1847  
 QY 1267 AACGTTGCTTCTCGAATAATAGAAAAACCTTTGTTTCTCAAAATGTTGCTTCTCTGTT 1326  
 Db 1848 AspLeuValPhe-----AsnLeuSerHisSerVal 1857  
 QY 1327 TACATTGCTTAAGCAACCAAGTGGACAGGCTTCTTAATACAGTAGTAATAATACGGA 1386  
 Db 1858 -----SerHisAlaIleAsnThrGlnGlyThr----- 1867  
 QY 1387 GATGAGGAGCTATCTTCTGTAAGAAATGCTGCGCAAGCAGGATCCAAATCTCGATCA 1446  
 Db 1868 -----AlaThrIleMetAlaAsnAsnAsnProLeu 1877  
 QY 1447 GTTTCCTTTGATGGAGGAGGAGTAGTTTCTTTTAGTACAAATGTTAGCTGCTGGGAAGG 1506  
 Db 1878 IleGlnPheAsn-----AlaSerSerLysGlu 1886



Db 2290 AlaAspProAsnLysLysSerValIleuThrIleAsnPheAlaLeuGluSerArgHis--- 2308

Qy 3574 GATGTTTACTTGAATGAGACTGGCGCTTTTGTCAATGCTCTATGCAGGAGCTCAAGAAGGT 3633

Db 2309 -----TyrPheAsnLysAsn---SerTyrTyrPheValIleAlaAspValGlyArgAsp 2325

Qy 3634 TCGATTATCTCAGCTAATGGCGACAAATTAAACGATTATCCGGACAAAACCATACATATTATCA 3693

Db 2326 LeuPheIleAsnSerMetGlyAspLysMetValArgPheIleGlyAsnAsnThrLeuSer 2345

Qy 3694 TTTACAGATTCTCAAGGCCAGTCTTCAAAATTATGCC---TTTCATTTCAGCAGGAGAG 3750

Db 2346 TyrArgAsp-----GlyGlyArgTyrAsnThrPheAlaSerIleIleThrGlyGlyGlu 2363

Qy 3751 ACATTACTCTGAGAGATTTTTCAGCTCGATGTTCTCGAAAAATGTTTCTTGGCGGAA 3810

Db 2364 -----IleArgLeuPheLysThr-----PheTyrValAsnAlaGlyIleGlyAla 2378

Qy 3811 AAGGAATGATCTCCGGGAAAACCGTGAGTATTTCCGGA 3849

Db 2379 ArgPheGlyLeuAspTyrLysAspIleAsnIleThrGly 2391

RESULT 34

US-11-052-554A-374

Sequence 374, Application US/11052554A

Publication No. US2005028866A1

GENERAL INFORMATION:

APPLICANT: Sachdeva, et al.

TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE

TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

FILE REFERENCE: 30853/40359A

CURRENT APPLICATION NUMBER: US/11/052,554A

CURRENT FILING DATE: 2005-02-07

PRIOR APPLICATION NUMBER: US 60/589,227

PRIOR FILING DATE: 2004-07-20

PRIOR APPLICATION NUMBER: IN 173/DEL/2004

PRIOR FILING DATE: 2004-02-06

NUMBER OF SEQ ID NOS: 763

SOFTWARE: PatentIn version 3.3

SEQ ID NO 374

LENGTH: 2015

TYPE: PRT

ORGANISM: Neisseria meningitidis Z2491

US-11-052-554A-374

Alignment Scores:

Pred. No.: 8,64e-09 Length: 2015

Score: 202.50 Matches: 298

Percent Similarity: 32.7% Conservative: 201

Best Local Similarity: 19.5% Mismatches: 585

Query Match: 2.6% Indels: 442

DB: 11 Gaps: 68

US-10-701-844-1 (1-4435) x US-11-052-554A-374 (1-2015)

Qy 448 GGGGGGGATATGCAGCAGAAATCATGTTCTCAAGGAATTTACGATGGGAGAGCTTA 507

Db 284 GlyGlyMetTyrAlaAspSerIle----- 291

Qy 508 ACTGTATCATTTCCCTACTACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTCTGCA 567

Db 292 -----ThrLeuIleAlaAsnGluLysGlyValGlyValLysAsnAla 305

Qy 568 GGAGAGTTAAACATTAAAAAATCTTGACAAATCTATTGACGCTTCCTGCTTAAAGTTGTTTT 627

Db 306 GlyThrLeu-----GluAlaAlaLysGlnLeuIleValThrSerSer 319

Qy 628 GGGAAGTTATTAGGAGT-----TTTACTGTTTTAGGAGAGGACACATCGTTG 675

Db 320 GlyArgIleGluAsnSerGlyArgIleAlaThrThrAlaAspGlyThrGluAlaSerPro 339

Qy 676 ACTTTCGAGAACATACGGACTTCTCAAAATGGGCGAGCT-----CTAAGTAATAGC 726

1702 QY -----GTAAGTGTCTCTCACAAGCCATTTCGATG 1731  
Db ThrAsnIleThrSerSerSerGlyAspIleThrLeuValAlaGlyAsnGlyIleGlnLeu 698  
1732 QY CGATCGGA-----GGGAAATAAGCACATTAAAGAGCTAAAGCA 1770  
Db GlyAspGlyLysGlnArgAsnSerIleAsnGlyLysHisIleSerIleLysAsnAsnGly 718  
1771 QY GGGCATCAGATTCCTTTAATGATCCCATCGAGATGGCAACGGA----- 1815  
Db GlyAsnAlaAspLeuLysAsnLeuAsnValHisAlaLysSerGlyAlaLeuAsnIleHis 738  
1816 QY -----AATAACCGACGCGCAGCTCTCC----- 1839  
Db SerAspArgAlaLeuSerIleGlnAsnThrLysLeuGluSerThrHisAsnThrHisLeu 758  
1840 QY -----AACTTCTAAATAATTAACGATGCTGAAGGATAC----- 1872  
Db AsnAlaGlnHisGluArgValThrLeuAsnGlnValAspAlaTyAlaHisArgHisLeu 778  
1873 QY -----ACAGGGGATATTGTTTT----- 1890  
Db SerIleThrGlySerGlnIleTrpGlnAsnAspLysLeuProSerAlaAsnLysLeuVal 798  
1891 QY GCTAATGGA-----AGCAGTACTTCTGACCAAAATGTTACGATA 1929  
Db AlaAsnGlyValLeuAlaLeuAsnAlaArgTySerGlnIleAlaAspAsnThrThrLeu 818  
1930 QY GAGCAAGGAAGGATTGTTCTTCGTAAGGCAAAATTA-----TCAGTGAAT 1977  
Db ArgAlaGlyAlaIleAsnLeuThrAlaGlyThrAlaLeuValLysArgGlyAsnIleAsn 838  
1978 QY -----TCTCTAAGTCAGACAGGTGG 1998  
Db TrpSerThrValSerThrLysThrLeuGluAspAsnAlaGluLysProLeuAlaGly 858  
1999 QY ACTCTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCCAACCAACCAACAG 2058  
Db ArgLeuAsnIleGluAlaGlySer-----GlyThrLeuThrIleGluPro 873  
2059 QY CCTCTCGCGCTAATCAGTTGATCAGCTTCCCAATCTGCATTTGTCTCTTTCTTTG 2118  
Db -----AlaAsnArgIleSerAlaHisThrAspLeuSerIleLysThrGlyLys 890  
2119 QY TTAGCAACAATCGAGTTACGAATCCTCTACAAATCTCCAGCGCAAGATTCAT 2175  
Db LeuLeuLeuSerAlaLysGlyGlyAsnAlaGlyAlaProSerAlaGlnValSerSerLeu 910  
2176 QY -----CCTGAGTCATTGTTAGCACA----- 2196  
Db GluAlaLysGlyAsnIleArgLeuValThrGlyLeuThrAspLeuArgLysSerIle 930  
2197 QY ACTGCTGGTCT-----GTTACAATTAGTGGGCTCTATCTTTTTCGAGGATTG 2244  
Db ThrAlaGlyLysAsnLeuValAlaThrThrLysGlyLysLeuAsnIleGluAlaVal 950  
2245 QY GATGATACAGCTATGATAGGTATGATGGTGTCT-----AATCAAAAATC 2295  
Db AsnAsnSerPheSerAsnTyPheProThrGlnLysAlaAlaGluLeuAsnGlnLysSer 970  
2296 QY AATGCTCTGAATTAACAGTTAGGAGTAAAGCCCAAGTAAATGCCCATCAGATTG 2352  
Db LysGluLeuGlnGlnIleAlaGlnLeuLysLysSerSerProLysSerLysLeuIle 990  
2353 QY -----ACTCTAGGGAATCAGATCCCTAAGTATGGCTATCAAGGAAGCTTGGGTG 2409  
Db ProThrLeuGlnGluArgAspArgLeuAlaPhe----- 1002  
2410 QY GATCCTTAATACGAATAATGCTCTTATCTCTGAAAGCTACATGGAACCTGGG 2469  
Db -----TyrIleGlnAlaIleAsnLysGluValLysGly 1013

2470 QY TATAATCCT---GGGCTGAGCGAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCC 2526  
Db LysLysProLysGlyLysGluTyThrLeuGlnAla-----LysLeuSerAlaGln 1029  
2527 QY ATTTTATATACATCTGCGCAT-----TCAGCAATTCAGCAAGT 2568  
Db AsnIleAspLeuIleSerAlaGlnGlyIleGluIleSerGlySerAspIleThrAlaSer 1049  
2569 QY GTGATGGGCTCTTATTGTCGAGGA----- 2595  
Db LysLysLeuAsnLeuHisAlaAlaGlyValLeuProLysAlaAlaAspSerGluAlaAla 1069  
2596 QY ---TTATGGGTTCTGGAGTTTCGAATTTCTTC----- 2625  
Db AlaIleLeuIleAspGlyIleThrAspGlnTyThrGluIleGlyLysProThrTyThrLysSer 1089  
2626 QY TATCATGACCGGATGCTTTAGGTACGGATATCGGTATATAGTGGG-----GGTTAT 2679  
Db HisTyThrAspLysAlaAlaLeuAsnLysProSerArg-----LeuThrGlyArgThrGlyVal 1108  
2680 QY TCCTTAGGAGCAAACTCTACTTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTA 2739  
Db SerIleHisAlaAlaAlaLeuAspAlaArgIleIleGlyAlaSerGluIle 1128  
2740 QY TTGCTAGATCTAAAGATTATGTAGTGTGCTTCCAAATCATCATGCTTGCATAGATCC 2799  
Db LysAlaProSer-----GlySer 1134  
2800 QY GTTTATCTATCTACCAACAACAGCTTTATGTGATCTTCTTCGGAGATCG----- 2853  
Db IleAspIleLysAlaHisSerAspIleValLeuGluAlaGlyGlnAsnAspAlaTyThr 1154  
2854 QY TTTATCCGCTGCTAGTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTGCA 2913  
Db PheLeuThrLysGlyLysSerGlyLysIleAlaArgLysThrLysPheThrSerThr 1174  
2914 QY GAGGAG-----AGCATGTTCTGGGATAATACTGCTGCTGCTGAGAG 2958  
Db ArgAspHisLeuIleMetProAlaProValGluLeuThrAlaAsnGlyIleThrLeuGln 1194  
2959 QY ATTGA-----GCGGATTACCGATT 2979  
Db AlaGlyGlyAsnIleGluAlaAsnThrThrArgPheAsnAlaProAlaGlyLysValThr 1214  
2980 QY GTGATTACTCCATCTAAGCTCTATTGAATGATGTTGCGTCTTCTGCAAGCTGAGTTT 3039  
Db LeuValAlaGlyGluGluLeuGlnLeu-----LeuAlaGluGlu 1227  
3040 QY TCTTATGCGGATCATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGCAATTCAGAGC 3099  
Db GlyIleHisLysHisGluLeu-----AspValGlnLysSerArgPhe----- 1242  
3100 QY GGACATCTCTAAATCTATCATGTTCTGTTGAGTGAAGTTTGTATCGATGT-----TCT 3153  
Db -----IleGlyIleLysValGlyLysSerAsnTyThr 1253  
3154 QY AGTACATCTCTAATAATATATAGCTTTATGCGCGCTTATATCTGTGATCTTATCGCACC 3213  
Db LysAsnGluLeuAsnGluThrLysLeuProValArgValAlaGlnThrAlaAlaThr 1273  
3214 QY ATCTCTGCTGCTGAGACACGCTCTCTATCCCATCAA----- 3249  
Db ArgSerGlyTyPheAspThrValLeuGluGlyThrGluPheLysThrThrLeuAlaGlyAla 1293  
3250 QY -----GAGCATGACACACAGATGCCCTTTCATTTCATTCAGACAT 3288  
Db AspIleGlnAlaGlyValGlyGlyLysAlaArgValAspAla-----LysIleLeuLys 1312  
3289 QY GAGTTGCTGTTAGGAGTCTATGATGCTTCTTAAACAAGTAATAGATATATAGC 3348  
Db GlyIleValAsnArgIleGlnSerGluGluLysLeuGluThrAsnSerThrValTrp----- 1331  
3349 QY CATGGAAGATATGATGATCGAGATGCTTCTTCGAGGCTATGTTTGTAGTGCAGGAAGTAGA 3408

```

1332 -----GlnLysGlnAlaGlyArgGlySerThrIleGluThrLeuLysLeu 1346
1340 GTCCGGTCTCAAAA -----TATTGGTTAGAT 3435
1347 ProSerPheGluSerProThrProLysLeuSerAlaProGlyGlyTyrIleValAsp 1366
1346 AGTTAAGTGTAGCGATGCTTTTCTTTCAGATCTACATCATTTGTTTGTAGCTTGT 3495
1367 -----IleProLysGlyAsnLeuLysThrGluIleGluLysLeuSerLys 1381
1396 TTGTGTTCTTATTCGTATGATTCGCGAGCTCTCTCAAGTGTAAAGCTTAATGTAAAC 3555
1382 GlnProGluTyrAlaTyr -----LeuLysGlnLeuGlnValAla 1394
1396 ACTCTCTTTAAGGAGCAGATGTTTACTTGAATGGAGAC ----- 3594
1395 LysAsnIleAsnTrpAsnGlnValGlnLeuAlaTyrAspArgTrpAspTyrLysGlnGlu 1414
1395 -----TGCGCTTTTGTCAATGTCATGTCAGGA 3621
1415 GlyLeuThrGluAlaGlyAlaAlaIleAlaLeuAlaValThrValThrSerGly 1434
1434 GCTGAAGAGTTCGATTTATCTACGTAATGCGCAATTTAACGATTACC ----- 3672
1435 AlaGlyThrGlyAlaValLeuGlyLeuAsnGlyAlaAlaAlaAlaThrAspAlaAla 1454
1435 -----GACAAACCATACATTATTCATTTACAGATTTCTCAAGGCGCAGTT--- 3717
1455 PheAlaSerLeuAlaSerGlnAlaSerValSerPheIleAsnAsnLysGlyAspValGly 1474
1474 -----CTTCAAAATTTATGCTTCATTTTCAGCA 3744
1475 LysThrLeuLysGluLeuGlyArgSerSerThrValLysAsn -----LeuValValAla 1492
1492 GGAGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAATATGTTCTTGC 3804
1493 AlaAlaThrAlaGlyValAlaAspLysIleGlyAlaSerAlaLeuAsnAsnValSer --- 1511
1511 GGAGAAAGGAATGATCTCCGGGAAACCTGAGTATTTCCGGAGCAGGCGAAGTGATT 3864
1512 ---AspLysGlnTrpIleAsnAsnLeuThrValAsnLeuAlaAsnAlaGlySerAlaAla 1530
1530 TTCTGGGATACTCCGTGGGTATTTCTCTTATCTACTGTGCGCAACCTCATCACT 3924
1531 LeuIleAsnThrAlaIleAsnGlyGlySerLeuLys ----- 1542
1542 CCGCTGCTCCACAGTTAGTATGCTCGGAAAGGCTCTATTTTCTGTAGAGACTAGT 3984
1543 ---AspAsnLeuGlyAspAlaAlaLeuGlyAlaIleValSerThrValHisGly 1559
1559 TTGGAGATCTCAGGCGTCAAAAAGGGTCTGTTTCGATAATTAATGCGGGAATTTCCGA 4044
1560 GluValAlaSerLysIleLysPheAsnLeuSerGluAspTyrIleThrHisLysIleAla 1579
1579 ACAGTTTTCAGGT -----AAGAATAATAATAATGCTGGTGGTGGAGGC 4089
1580 HisAlaIleAlaGlyCysAlaAlaAlaAlaAlaAlaLysGlyLysCysGlnAspGly-Al 1599
1599 AGTGGGTTCCGCTACA -----CCATCAAGTAC 4116
1599 aIleGlyAlaAlaValGlyGluIleValGlyGluAlaLeuThrAsnGlyLysAsnProAl 1619
1619 GACTTTTACAGTTAAACCTGTAAGGGAAGTTCTTTCACAGATAACGTAGCTCTTG 4176
1619 aThrLeuThrAlaLysGluArgGluGlnIleLeuAlaTyrSerLysLeuValAlaGlyTh 1639
1639 CGAGGCGGAGTGGTT 4192
1639 rValSerGlyValVal 1644

```

RESULT 35

US-11-045-208-38

```

; Sequence 38, Application US/11045208
; Publication No. US20060063167A1
; GENERAL INFORMATION:
; APPLICANT: (A) NAME : I.N.S.E.R.M
; TITLE OF INVENTION: DNA, specific proteins and peptides
; of the Neisseria meningitidis species bacteria, method
; for obtaining them and their biological application.
;
; NUMBER OF SEQUENCES: 99
; STREET: New Horizons Court
; CITY: Brentford
; COUNTRY: Great-Britain
; ZIP: TW 89 EP
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OBB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/045,208
; FILING DATE: 31-Jan-2005
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/214,759
; FILING DATE: 22-APR-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: SMITHKLINE BEECHAM
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1981 acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..1981
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
;
; US-11-045-208-38
;
; Alignment Scores:
; Pred. No.: 1,73e-08 Length: 1981
; Score: 199.00 Matches: 306
; Percent Similarity: 32.5% Conservative: 200
; Best Local Similarity: 19.6% Mismatches: 591
; Query Match: 2.5% Indels: 461
; DB: Gaps: 71
;
; US-10-701-844-1 (1-4435) x US-11-045-208-38 (1-1981)
;
QY 448 GGGGGGGGATATGAGCGAAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAGACGTTA 507
DB 284 GlyGlyMetTyrAlaAspSerIle----- 291
QY 508 ACTGTATCATTTCCCTATACCTACTGTTATAGAGATCCGAGTGGGACTACTGTTTCTGCA 567
DB 292 -----ThrLeuIleAlaAsnGlnLysGlyValGlyValLysAsnAla 305
QY 568 GGAGAGTTAACAATAAAAAATCTTGACAAATCTTATGCGAGCTTTGCTTAAAGTTGTTT 627
DB 306 GlyThrLeu-----GluAlaAlaLysGlnLeuIleValThrSerSer 319
QY 628 GGGAACTATTATGGAGAT-----TTTACTGTTTGGGAGAGGACACTCGTTG 675
DB 320 GlyArgIleGluAsnSerGlyArgIleAlaThrAlaAspGlyThrGluAlaSerPro 339
QY 676 ACTTTCGAGAACATACGAGCTTCTACAAATCGGGCAGCT-----CTAAGTAATAGC 726
DB 340 ThrThrLeuSerIleGluThrThrGlnLysGlyAlaAlaGlyThrPheIleSerAsnGly 359
QY 727 GCT-----GCTGATGGACTGTTTACTATTGAGGTTTAAAGAAATATCTCTTTCC 777
DB 360 GlyArgIleGluSerLysGlyLeuValIleGluThrGlyGluAspIleSerLeuArg 379
QY 778 AATGCAATTCATTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGATGACCGACT 837

```



QY 2569 GTGGATGGCGCTCTATTATGCGAGG----- 2595  
 Db 1050 LysLysLeuAsnLeuHieAlaAlaGlyValLeuProLysAlaAlaAspSerGluAlaAla 1069  
 QY 2596 ---TTATGGGTTCTGGAGTTTCGAATTCCTC----- 2625  
 Db 1070 AlaIleLeuIleAspGlyIleThrAspGlnIleGlyLysProThrTyrLysSer 1089  
 QY 2626 TATCATGACCGGATGCTTTAGGTACGGATATCGTATATTAGTGGG-----GGTTAT 2679  
 Db 1090 HieTyrAspLysAlaAlaLeuAsnLysProSerArg---LeuThrGlyArgThrGlyVal 1108  
 QY 2680 TCCTTAGAGCAATCCTCTACTTTGGATCATCGATGTTTGGTCTAGCATTTACCAAGTA 2739  
 Db 1109 SerIleHieAlaAlaAlaLeuAspAlaArgIleIleGlyAlaSerGluIle 1128  
 QY 2740 TTTGGTAGATCTAAAGATTATGTAGTGTCTGCTTCCATCATCATGCTTGATAGATCC 2799  
 Db 1129 LysAlaProSer-----GlySer 1134  
 QY 2800 GTTTATCTCTACCAACAGCTTTATGTGATCCTATTGTTGCGAGATGCG----- 2853  
 Db 1135 IleAspIleLysAlaHieSerAspIleValLeuGluAlaGlyGlnAsnAspAlaTyrThr 1154  
 QY 2854 TTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCA 2913  
 Db 1155 PheLeuLysThrLysGlyLysSerGlyLysIleIleArgLysThrLysPheThrSerThr 1174  
 QY 2914 GAGGAG-----AGCATGTTCTGTTGGGATAATACTGCTGGCTCGAGAG 2958  
 Db 1175 ArgAspHieLeuIleMetProAlaProValGluLeuThrAlaAsnGlyIleThrLeuGln 1194  
 QY 2959 ATTGGA-----GCGGGATTACCGATT 2979  
 Db 1195 AlaGlyLysAsnIleGluAlaAsnThrThrArgPheAsnAlaProIleGlyLysValThr 1214  
 QY 2980 GTGATTACTCCATCTAAGCTCTATTGTAATGATGTTGCGTCTTTCGTGCAAGCTGAGTTT 3039  
 Db 1215 LeuValAlaGlyGluGluLeuGlnLeu-----LeuAlaGluGlu 1227  
 QY 3040 TCTTATGCGGATCATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCGATTCAGAGC 3099  
 Db 1228 GlyIleHieLysHieGluLeu-----AspValGlnLysSerArgPhe----- 1242  
 QY 3100 GGACATCTCTAAATCTATCAGTCTCTGTTGAGTCAAGTTTGTATCGATGT-----TCT 3153  
 Db 1243 -----IleGlyIleLysValGlyLysSerAsnTyrSer 1253  
 QY 3154 AGTACACATCTTAATAATATAGCTTTATGCGGGCTTATATCTGTGATGCTTATCGCAC 3213  
 Db 1254 LysAsnGluLeuAsnGluThrLysLeuProValArgValValAlaGlnThrAlaAlaThr 1273  
 QY 3214 ATCTCTGTTACTGAGCAAGCTCTCTATCCATCA----- 3249  
 Db 1274 ArgSerGlyTyrAspThrValLeuGluGlyThrGluPheLysThrThrLeuAlaGlyAla 1293  
 QY 3250 -----GAGACATGGACACAGATGCTTCATTTAGCAAGACAT 3288  
 Db 1294 AspIleGlnAlaGlyValGlyGluLysAlaArgValAspAla---LysIleIleLeuLys 1312  
 QY 3289 GGAGTTGTGGTTAGAGATCTATGATGCTTCTTAACAAGTAATATAGATATATGGC 3348  
 Db 1313 GlyIleValAsnArgIleGlnSerGluGluLysLeuGluThrAsnSerThrValTyr--- 1331  
 QY 3349 CATGGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGGTTTGTAGTCAGGAAGTAGA 3408  
 Db 1332 -----GlnLysGlnAlaGlyArgGlySerThrIleGluThrLeuLysLeu 1346  
 QY 3409 GTCCGGTCTAAAA-----TATTGGTTAGAT 3435  
 Db 1347 ProSerPheGluSerProThrProProLysLeuSerAlaProGlyGlyTyrIleValAsp 1366

QY 3436 AGTTAAGTGTAGCGATGCGCTTTTCTTTGAGATCTACATCATTTTGTTTTGTAGCTGT 3495  
 Db 1367 -----IleProLysGlyAsnLeuLysThrGluIleGluLysLeuSerLys 1381  
 QY 3496 TTGTGTTCTTATCGTATGATTCGCGAGCTCTCTCAAGTGTAAACCCCTAATAACCC 3555  
 Db 1382 GlnProGluTyrAlaTyr-----LeuLysGlnLeuGlnValAla 1394  
 QY 3556 ACTCCTTTTAAAGGAGAGCATGTTTACTTTGAATGGAGAC----- 3594  
 Db 1395 LysAsnIleAsnTyrAspArgTyrAspArgTyrAspTyrLysGlnGlu 1414  
 QY 3595 -----TGCGCTTTTGTCAATGCTCATGCAGA 3621  
 Db 1415 GlyLeuThrGluAlaGlyAlaAlaIleIleAlaLeuAlaValThrValThrSerGly 1434  
 QY 3622 GCTGAAGAGGTTTCGATTATCTCACTAATCGGACCAATTTAAACGATPACC----- 3672  
 Db 1435 AlaGlyThrGlyAlaValLeuGlyLeuAsnGlyAlaAlaAlaAlaThrAspAlaAla 1454  
 QY 3673 -----GGCAAAACCATACATTATTTACAGATTCTCAAGGGCCAGTTCT 3720  
 Db 1455 PheAlaSerLeuAlaSerGlnAlaSerValSerPheIleAsnAsnLysGlyAspVal--- 1473  
 QY 3721 CAAATTTATGCTTCATTTTCAGCAGGAGACACTTACTCTGAGAGATTTTTCGAGTCTG 3780  
 Db 1474 -----GlyLysThrLeuLys----- 1478  
 QY 3781 ATGTTCTCGAAAAATGTTTCTTGGGAGAAAGGAATGATCTCCGGGAAACCCGTGAGT 3840  
 Db 1479 -----GluLeuGlyArgSerSerThrValLysAsnLeuValValAla 1492  
 QY 3841 ATTTCCGAGCAGCGAAGTGAATTTTCGGATTAATCTCCGGGGGTATTCCTTTTATCT 3900  
 Db 1493 AlaAlaThrAlaGlyValAla-----AspLysIleGlyAlaSerAlaLeuAsn 1508  
 QY 3901 ACTGTGCCAATCATCATCACTCACTCCGCTCTCCACAGATTAGTGTCCGGAAGGG 3960  
 Db 1509 AsnVal---SerAspLysGlnTyrIleAsnAsnLeuThrValAsnLeuAlaAsnAlaGly 1527  
 QY 3961 TCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAAAAAGGGTCTATGTTTC 4020  
 Db 1528 SerAlaAlaLeuIleAsnThrAlaVal----- 1536  
 QY 4021 GATAATAATCGCGAATTTCCGAAACAGTCTTTTCGAGTAAGATAATAATAATCTCGGT 4080  
 Db 1537 -----AsnGlyGlySerLeu-LysAspAsnLeuGluAlaAsnIleLeuAlaAlaLeuVa 1554  
 QY 4081 GGTGGAGGAGTGGTTCGCTTACCATCACCATCAAGTACGACTTTTACAGTTAAAACTGTAA 4140  
 Db 1554 AsnThrAlaHieGlyGluAlaAlaSerLysIleLysGlnLeuAspGlnHieTyrIleVa 1574  
 QY 4141 AGGGAAGTGTCTTTCACAGATAAGTACGCTCTTTCGCGAGCGCGAGTGGTTTATAAAGG 4200  
 Db 1574 LhisLysIleAlaHieAla-----IleAlaGlyCysAlaAlaAlaAlaAsnLysGly 1592  
 QY 4201 CATTTGTCTTTTCAAGACAATGAAGGAGGCGATA----- 4234  
 Db 1592 y-----LysCysGlnAspGlyAlaIleGlyAlaAlaValGlyGluIleValGlu 1608  
 QY 4235 -----TCTTCCGAGGGAAC-----ACGCATACGATGATTTAAGGAT 4272  
 Db 1608 yGluAlaLeuThrAsnGlyLysAsnProAspThrLeuThrAlaLysGluArgGluGlnIle 1628  
 QY 4273 TCTTCTCTCTACTACTCAGGATCAGATAACGAGACAGGAGCGGTGGAGGA 4324  
 Db 1628 eLeuAlaTyrSerLysLeuValAlaGlyThrValSerGlyValValGlyGly 1645

RESULT 36

US-11-067-260-20

; Sequence 20, Application US/11067260

; Publication No. US20060051840A1

; GENERAL INFORMATION:



```
; APPLICANT: Arico, Maria
; TITLE OF INVENTION: HETEROLOGOUS EXPRESSION OF NEISSERIAL PROTEINS
; FILE REFERENCE: 223002099501
; CURRENT APPLICATION NUMBER: US/11/067,260
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: 10/220,480
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: PCT/IB01/00420
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1312
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: deletion G983-741
US-11-067-260-20

Alignment Scores:
Pred. No.: 2,37e-08 Length: 1312
Score: 196.50 Matches: 250
Percent Similarity: 31.3% Conservative: 151
Best Local Similarity: 19.5% Mismatches: 441
Query Match: 2.5% Indels: 439
DB: 11 Gaps: 62

US-10-701-844-1 (1-4435) x US-11-067-260-20 (1-1312)
QY 559 TTTTCTGCGAGGAGTTAACTTAAAAATCTTGACAAATCTATTGCGAGCTTTGCGCTTTA 618
DB 268 TyrSerGlyAspLysThr-----AspGluGlyIleArgLeuMetGln 283
QY 619 AGTTGTTTGGGAACCTATTAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACT 678
DB 284 SerAspTyrGlyAsnLeu-----SerTyrHisIleArgAsnLysAsnMetLeuPheIle 301
QY 679 TTC-----GAGAACATACCGACTTCTACAAATGGCGCAGCTCTA----- 717
DB 302 PheSerThrGlyAsnAlaGlnAlaGlnProAsnThrTyrAlaLeuLeuProPheTyr 321
QY 718 AGTAATAGCGCTGCTGATGACTGTTTACTATTAGAGGT----- 756
DB 322 GluLysAspAlaGlnLysGlyIleIleThrValAlaGlyValAspArgSerGlyGluLys 341
QY 757 TTTAAAGAAATTATCTCTT-----TCCAAATGTC 783
DB 342 PheLysArgGluMetTyrGlyGluProGlyThrGluProLeuGluTyrGlySerAsnHis 361
QY 784 AATTCATTCTGCGTACTGCTGCTGCGAACGACTAATAGGTAGCCAGACTCCGACG 843
DB 362 CysGlyIleThrAlaMetTyrCysLeuSerAlaProTyrGluAlaSerValArgPheThr 381
QY 844 ACAACACT-----ACACCGCTAATGCTACTATTATTCTTAAACA 885
DB 382 ArgThrAsnProIleGlnIleAlaGlyThrSerPheSerAlaProIleValThrGlyThr 401
QY 886 GATCTTTTGTACTCAATAATAGAGATTCTCATTTCTATAGTAATTTAGTCTCTGGAGAT 945
DB 402 AlaAlaLeuLeuLeu-----GlnLysTyrProTyrMetSerAsn----- 414
QY 946 GGGGAGCTATAGATGCTAAGAGCTTAAACGCTTCAAGGTTCAAGAAATTACAGCTTTGTCTTC 1005
DB 415 -----AspAsnLeuArgThrThr 420
QY 1006 CAAGAAATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCAACAGTTTC----- 1059
DB 421 LeuLeuThrThrAlaGln-----AspIleGlyAlaValGlyValAspSerLysPheGlyTyr 439
QY 1060 -----TCTGCTATGGCTAACGAGGCTCTATTGCTTTGTAGCGCAAT 1101
DB 440 GlyLeuLeuAspAlaGlyLysAlaMetAsnGlyProAlaSerPheProPheGlyAspPhe 459
```

```
QY 1102 GTTCGAGGTAAGAGGGGAGGAGATTGCTGCTGTTCCAGATGGCGACAGGAGTGTC 1161
DB 460 ThrAlaAspThrLysGly----- 465
QY 1162 TCATCTACTTCAACAGAAGATCCAGTAGTAGTATTTTCCAGAAATACTGCGGTAGATT 1221
DB 466 -----ThrSerAspIleAlaTyrSerPhe---ArgAsnAspIleSerGlyThr 480
QY 1222 GATGGAACTAGCCCGAGTAGGAGGAGGATTACTCTCTCGGAACTGCTCTTCTTC--- 1278
DB 481 GlyGlyLeuIleLysGlyGlySerGlnLeuGlnHisGlyAsnAsnThrTyrThr 500
QY 1279 -----CTGAATAATGAAAAACCTTTGTTTCTCAACAATGTTGCTCTCTCTGTT 1326
DB 501 GlyLysThrIleIleGluGlySerLeuValLeuTyrGlyAsnAsnLysSerAspMet 520
QY 1327 TACATTGCTGCTAAGCAACCAACA-----AGTGACAGGCTCTTAATACGAGTAATA 1380
DB 521 ArgValGluThrLysGlyAlaLeuIleTyrAsnGlyAlaAspSerGlyGlySerLeuAsn 540
QY 1381 TACGGAGATGGAGGAGCTATCTTCTGTAAGAATGGTGGCGACAGGATCCAAATACTCT 1440
DB 541 -----SerAspGlyIleValTyrLeuAlaAspThrAspGlnSerGlyAlaAsnGluThr 558
QY 1441 -----GGATCAGTTTCTTGTGATGAGAGGAGTAGTATTTTCTTTAGTACCAAT 1488
DB 559 ValHisIleLysGlySerLeuGlnLeuAspGlyLysGlyThrLeuTyr----- 574
QY 1489 GTAGTGTCTGGAAA-----CGGCGAGCTATTTAT 1518
DB 575 ThrArgLeuGlyLysLeuLysValAspGlyThrAlaIleIleGlyGlyLysLeuTyr 594
QY 1519 GCCAAAAG----- 1527
DB 595 MetSerAlaArgGlyLysGlyAlaGlyTyrLeuAsnSerThrClyArgArgValProPhe 614
QY 1528 CTCTCGGTTGCTAACTGTGCGC---CCTGTACAAATTTTAAAGAAATATCGCTAATAGTGT 1584
DB 615 LeuSerAlaAlaLysIleGlyGlnAspTyrSerPhePheThrAsnIleGluThrAspGly 634
QY 1585 GGAGCGAATTTTATGAGAAATCTGGAGACTCGATTTTATCTGCTGATATGAGAT--- 1641
DB 635 GlyLeuLeuAlaSerLeuAspSerValGluLysThrAlaGlySerGlu---GlyAspThr 653
QY 1642 -----ATTATTTTCGATGGAATCTTAAAGAACACGCCAAA----- 1677
DB 654 LeuSerTyrTyrValArgArgGlyAsnAlaAlaArgThrAlaSerAlaAlaHisSer 673
QY 1678 -----GAGAATGCTGCGGATGTTTAAATGGCGTAACTGTGTCTCTCACAAGCC 1722
DB 674 AlaProAlaGlyLeuLysHisAlaValGluGlnGlyGlySerAsnLeuGluAsnLeuMet 693
QY 1723 ATTTCCAGTGGGA-----TCGGGAGGAAATACGCATTAAGACATTAAGACGAGG 1773
DB 694 ValGluLeuAspAlaSerGluSerSerAlaThrProGluThrValGluThrAlaAla 713
QY 1774 CATCAGATTCTTTAATGATCCCATCGAGATGCGAAACGAAATACCAAGCCAGCCAG 1833
DB 714 -----AspArgThrAspMetProGlyIleArgProTyrGlyAlaThr 727
QY 1834 TCTTCCAACTTCTAAAAATTAACGATGGTGAAGGATACACAGGGGATATTTGTTTGTCT 1893
DB 728 PheArgAlaAlaAlaValGlnHisAlaAsnAlaAlaAspGlyValArgIlePheAsn 747
QY 1894 AATGGAAGCAGTACTTTGTACCAAAATGTTTACGATAGAG-----CAAGGAAGG 1941
DB 748 SerLeuAlaAlaThrValTyrAlaAspSerThrAlaAlaHisAlaAspMetGlnGlyArg 767
QY 1942 ATTGTTCTTCTGTAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTCGAGT 2001
DB 768 -----ArgLeuLysAlaValAspGlyLeuAspHisAsnGlyThrGly 782
QY 2002 CTGTATATGGAAGCT-----GGGAGTACATGGGAT----- 2031
```

Db	783	LeuArgValIleAlaGlnThrGlnGlnAspGlyGlyThrTrpGluGlnGlyGlyValGlu	802
QY	2032	-----TTGTAACTCCACCAACCACCAACAGCCT	2061
Db	803	GlyLysMetArgGlySerThrGlnThrValGlyIleAlaAlaIysThrGlyGluAsnThr	822
QY	2062	CTGCGCCCTAATCAGTTGATCAGCGCTTCCAAATCTGCATTTGTCTCTTCTCTCTTGTGTTA	2121
Db	823	ThrAlaAla-----AlaThrLeuGlyMetGlyArgSerThrTrpSer	836
QY	2122	GCAACAATCAGATTAGCAATCTCTTACCAATCTCCACCGCAAGATTTCTCATCTCGCA	2181
Db	837	GluAsnSerAla-----AsnAlaLysThr-----AspSerIleSerLeu	849
QY	2182	GTCAATTGGT---AGCACAACTGCTGGTCTGTTACAATTAGTAGGGCTATCTTTTGTAG	2238
Db	850	PheAlaGlyIleArgHisAspAlaGlyAspIleGlyTyrLeuLysGlyLeuPhe-----	867
QY	2239	GAATTGGATGATACAGCTTATGATAGTATGAT-----TGGCTAGGTTCT	2283
Db	868	-----SerTyrGlyArgTyrLysAsnSerIleSerArgSerThrGlyAla	882
QY	2284	AATCAA-----AAAATCAATGTCCTGAAATTACAGTTAGGACTAAGCCCCCA	2331
Db	883	AspGluHisAlaGluGlySerValAsnGlyThrLeuMetGlnLeuGlyAlaLeuGlyGly	902
QY	2332	GCTAATGCCCA-----TCGATTTCAGTCTTAGGGAT-----	2364
Db	903	ValAsnValProPheAlaAlaThrGlyAspLeuThrValGluGlyLeuArgTyrAsp	922
QY	2365	-----QAGATGCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCCTAAT	2418
Db	923	LeuLeuLysGlnAspAlaPheAlaGluLysGlySer---AlaLeuGlyTrpSerGlyAsn	941
QY	2419	ACAGCAATAATGTCCTTACTCTGAAAGCTACATGGAATAAACTGGGTATATCTCT	2478
Db	942	SerLeuThrGluGly-----	946
QY	2479	GGGCTGAGCGAGTAGCTCTTTGGTTCCAAATAGTTATGGGATCCATTTTAGAT---	2535
Db	947	-----ThrLeuValGlyLeuAlaGlyLeuLysLeuSerGlnProLeuSerAspLys	963
QY	2536	---ATCGATCTGGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTGTGCA	2592
Db	964	AlaValLeuPheAlaThrAlaGlyValGluArgAspLeuAsnGlyArgAspTyrThr---	982
QY	2593	GGATTATGGGTTCTGGAGTTTCGAATTTCTTCTATCATCAACCGCGATCTTTAGGTCAG	2652
Db	983	-----ValThrGlyGlyPheThrGlyAlaThrAlaAlaThrGlyLys	996
QY	2653	-----GGATATCGGTATATTAGTGGGGTTATTCTTTAGGCA	2691
Db	997	ThrGlyAlaArgAsnMetProHisThrArgLeuValAlaGly-----LeuGlyAla	1013
QY	2692	AATCTTACTTTGGATCATCGATGTTTGGTCTAGCAATTTACCGAAGTATTGGTAGATCT	2751
Db	1014	AspValGluPheGlyAsnGlyTrpAsnGlyLeuAlaArgTyrSerTyrAlaGly---Ser	1032
QY	2752	AAAGATTATGTAGTGTGCTTCCAAATCATCANGCTTGGCATAGGATCCGTTTATCTATCT	2811
Db	1033	LysGlnTyr-----GlyAsnHisSerGlyArgValGly-----	1043
QY	2812	ACCCAAACAGCTTTATGTGGATCCTATTTGTTCCGAGATGGGTTTATCCGTGTAGCTAC	2871
Db	1044	-----ValGlyTyrArgPheLeuGluGlySerGly	1053
QY	2872	GGGTTTGGGAATCAGCATATGAAAACTCATATACATTTTGCAGAGGAGCGATGTTCGT	2931
Db	1054	GlyGlyGly-----	1056
QY	2932	TGGGATATAAATCTCTCTGGCTGGAGAGATTGGACGGGATTCACGATTGTGATTACTCCA	2991

1057	Db	-----ValAlaLaAAspIleGlyAlaGlyLeuAlaAspAlaLeuThrAla	1071
2992	QY	TCTAAGCTCTATTGTAAGTAGTTGCGCTCTTCGTGCAAGCTGAGTTTCTTATGCGCAT	3051
1072	Db	ProLeuAspHisLysAspLysGlyLeuGlnSerLeuThrLeuAspGlnSerValArgLys	1091
3052	QY	CATGAATCTTTTACAGAGGAAGCGATCAAGCT---CGGGCATTCAAGAGCGACATCTC	3108
1092	Db	AsnGluLysLeuLysLeuAlaLaGlnGlyAlaGluLysThrTyrgLyAsnGlyAspSer	1111
3109	QY	CTAATCTTATCAGTCTCTGTTGCGAGTGAAGTTCATCGATGTTCTAGTACACATCTCTAAT	3168
1112	Db	LeuAsnThrGly-----LysLeuLysAsnAspLysValSer-----	1123
3169	QY	AAATATAGCTTTATGGCGCTTATATATCTGTGATGCT-----	3204
1124	Db	ArgPheAspPheIleArgGlnIleGluValAspGlyGlnLeuIleThrLeuGluSerGly	1143
3205	QY	---TATCGCACCATCTCTGTACTGTGACAGACAACGCTCTATCCCATCAAGAGACATGGACA	3261
1144	Db	GluPheGlnValTyrlsGlnSerHisSerAlaLeuThrAlaPheGlnThrGluGlnIle	1163
3262	QY	ACAGATCGCTTTCATTTAGCAAGACATGGAGTGTGTGGTAGA-----GGA	3306
1164	Db	GlnAspSerGluHisSerGlyLys-----MetValAlaLysArgGlnPheArgIleGly	1181
3307	QY	TCTATGATGCTTCTTAACAAGTAATATAGAAGTATATATGCCCATGGAAGATATGAGTAT	3366
1182	Db	AspIleAlaGlyGluHisThrSerPheAspLysLeuProGluGlyArgAlaThrTyr	1201
3367	QY	CGAGATCGTCTCGAGGCTATGTTGTGAGTGCAGGAAGTAGAGTCCGGTCTTAAATAATAT	3426
1202	Db	ArgGlyThrAlaPheGlySerAsp---AspAlaGlyGlyLysLeu-----	1215
3427	QY	TGGTTAGATAGTTAAGTGTTAGCGATGCGCTTTTCTTTGGAGATCTACATCTTTGTTTT	3486
1215	Db	-----	1215
3487	QY	TTAGCTGTTGTGTCTTCTATTTCGTATGGATTTCGCGAGC-----	3525
1216	Db	-----ThrTyThrIleAspPheAlaAlaLysGlnGlyAsnGlyLysIle	1230
3526	QY	---TCTCTCTCAAGTGTAAACGCTTAATGTAAACCACCTCCTTTTAAAGGGAGAC	3573
1231	Db	GluHisLeuLysSerProGlu-----	1237
3574	QY	GATGTTTACTTGAATGGAGACTGCGCTTTTGTCAATGCTCTATGCGAGAGCTGAAGAGGT	3633
1238	Db	---LeuAsnValAspLeuAlaAlaAspIleLysProAspGlyLysArgHis	1254
3634	QY	TGATATTCTCAGCTAATGGCGCAATTTAAACGATTACCGGACAAACCACATACATTATCA	3693
1255	Db	AlaValIleSer-----	1258
3694	QY	TTTACAGATTCTCAAGGCCAGTTCTTCAAAATTTATGCTTCATTTTCAGCAGGAGAGACA	3753
1259	Db	---GlySerValLeuTyraAsnGlnAla-----	1266
3754	QY	CTTACTCTGAGAGATTTTTCGAGTCTCATGTTCTCGAAAAATGTTTCTTCGGGAGAAAAG	3813
1267	Db	-----GluLys	1268
3814	QY	GGAAATGATCTCC-----GGGAAACCGTGAATATTTCCGGAGCAGCGCAA	3858
1269	Db	GlySerTySerLeuGlyIlePheGlyGlyLysAlaGlnGluValAlaGlySerAlaGlu	1288
3859	QY	GTG	3861
1289	Db	Val	1289

RESULT 37  
US-11-067-260-18  
; Sequence 18, Application US/11067260



```
QY 2002 CTGTATATGGAAGCT-----GGGAGTACATGGAT----- 2031
Db 783 LeuArgValIleAlaGlnThrGlnGlnAspGlyGlyThrGluGlnGlyValGlu 802
QY 2032 -----TTTGTAACCTCCACACACACAGAGCT 2061
Db 803 GlyLysMetArgGlySerThrGlnThrValGlyIleAlaAlaLysThrGlyGluAsnThr 822
QY 2062 CTGCGCGCTAATCAGTTGATCAGCTTCCCAATCTGCATTTGTCTTCTTCTTGTGTA 2121
Db 823 ThrAlaAla-----AlaThrLeuGlyMetGlyArgSerThrTrpSer 836
QY 2122 GCAACAATGAGTACGAATCTCTACCAATCTCCAGCGCAAGATTCATCTCTGCA 2181
Db 837 GluAsnSerAla-----AsnAlaLysThr-----AspSerIleSerLeu 849
QY 2182 GTCATTGGT---AGCACAACTGCTGTTCTGTTCACAAATAGTGGGCTTATCTTTTTCAG 2238
Db 850 PheAlaGlyIleArgHisAspAlaGlyAspIleGlyTyrLeuLysGlyLeuPhe----- 867
QY 2239 GATTGGATGATACAGCTTATGATAGTATGAT-----TGCGTAGGTTCT 2283
Db 868 -----SerTyrGlyArgTyrLysAsnSerIleSerArgSerThrGlyAla 882
QY 2284 AATCAA-----AAATCAATGTCCTGAAATTTACAGTTAGGACTAAGCCGCCA 2331
Db 883 AspGluHisAlaGluGlySerValAsnGlyThrLeuMetGlnLeuGlyAlaLeuGlyGly 902
QY 2332 GCTAATGCCCA-----TCAGATTGACTCTAGGAAT----- 2364
Db 903 ValAsnValProPheAlaAlaThrGlyAspLeuThrValGluGlyGlyLeuArgTyrAsp 922
QY 2365 -----GAGATGCCCTAAGTATCGCTCAAGGAAGCTTGGCGTGGGATCCTAAT 2418
Db 923 LeuLeuLysGlnAspAlaPheAlaGluLysGlySer---AlaLeuGlyTrpSerGlyAsn 941
QY 2419 ACAGCAATAATAGGT-----CCTATATCT 2442
Db 942 SerLeuThrGluGlyThrLeuValGlyLeuAlaGlyLeuLysLeuSerGlnProLeuSer 961
QY 2443 CTGAAGCTACATGACTAAACTGGGTATATCTGGCGCTGAGCGAGTAGCTCTTTTG 2502
Db 962 AspLysAlaValLeuPheAlaThrAla-----GlyValGluArgAspLeuAsnGly 978
QY 2503 GTTCCAAATAGTTATGGGGATCCATTTTATAGATACGATCTGCG----- 2547
Db 979 ArgAspTyrThrValThrGlyGlyPheThrGlyAlaThrAlaAlaThrGlyLysThrGly 998
QY 2548 -----CATTCAGCAATTCAGCAAGTGGAGTGGCGGCTCTTATTTGTGCA 2592
Db 999 AlaArgAsnMetProHisThrArgLeuValAlaGlyLeuGlyAlaAspValGluPheGly 1018
QY 2593 GGAATTATGGTTCTCGAGTTTCGAATTTCTCTAT-----CAT 2631
Db 1019 AsnGlyTrp---AsnGlyLeuAlaArgTyrSerTyrAlaGlySerLysGlnTyrGlyAsn 1037
QY 2632 GACCGCGATGCTTTAGTCAGGATATCGGTATATTAGTGGGGGTATTCTCTAGGA--- 2688
Db 1038 HisSerGlyArgValGlyValGlyTyrArgPheLeuAspGlyGlyGlyThrGlySer 1057
QY 2689 -----GCAACTCTCTATTGGATCATCGATGTTGGTCTAGCATTTACCAAGTA 2739
Db 1058 SerAspLeuAlaAsnAspSerPheIleArgGlnValLeuAspArgGlnHisPheGluPro 1077
QY 2740 TTTGCTAGA-----NCTAAAGATTATGATGTGCTGCTTCCCAATCATCAT 2784
Db 1078 AspGlyLysTyrHisLeuPheGlySerArgGlyGluLeuAlaGluArgSerGlyHis--- 1096
QY 2785 GCTTGATAGGATCCGTTTATCTATCTACCAAGCTTTTATGTGGATCTTATTTGTC 2844
Db 1097 -----IleGlyLeuGlyLysIleGlnSerHisGln-----LeuGlyAsnLeuMetIle 1112
```

```
QY 2845 GGAGATGCTTTATCTCCGTGCTAGCTACGGGTTT-----GGGAAT 2883
Db 1113 GlnGlnAlaAlaIleLysGlyAsnIleGlyTyrIleValArgPheSerAspHisGlyHis 1132
QY 2884 CAGCATATGAAAACCTCATATACATTTGCGAGGAGAGCGAT----- 2925
Db 1133 GluValHisSerProPheAspAsnHisAlaSerHisSerAspSerAspGluAlaGlySer 1152
QY 2926 -----GTTCTGGGATATAAT----- 2940
Db 1153 ProValAspGlyPheSerLeuTyrArgIleHisTrpAspGlyTyrGluHisPheProAla 1172
QY 2941 AACTGTCTGGCTGAGAGATTTGGAGCGGATTAACGATTGCGATTGATTACTCCATTAAGCTC 3000
Db 1173 AspGlyTyrAspGlyProGlnGlyGlyTyrProAlaProLysGlyAlaArgAspIle 1192
QY 3001 TATTGGAATGAGTTGCGTCCTTCTCGTCAAA---GCTGAGTTTCTTATGCGCATCATGAA 3057
Db 1193 TyrSerTyrAspIleLysGlyValAlaGlnAsnIleArgLeuAsnLeuThrAspAsnArg 1212
QY 3058 TCTTTTACAGAGGAGGCGATCAAGCTCGGCATTCAGAGCGGACATCTCTCTAAATCTA 3117
Db 1213 Ser---ThrGlyGlnArgLeuAlaAspArgPheHisAsnAlaGlySerMetLeuThrGln 1231
QY 3118 TCAGTTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACATCTCTAAATAATATAGC 3177
Db 1232 GlyValGlyAspGlyPheLys-----ArgAlaThrArgTyrSer 1244
QY 3178 TTTATGGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGTGTTACTGAGACACGCTC 3237
Db 1245 ProGluLeuAspArgSerGlyAsnAlaAlaGluAlaPheAsnGlyThr----- 1260
QY 3238 CTATCCCATCAAGAGACATGAGCAACAGATGCCCTTTTCATTTAGCAAGACATGGAGTTGTG 3297
Db 1260 ----- 1260
QY 3298 GTTAGAGGATCTATGTTATGCTTCTTAACAGTAATATA-----GAAGTA 3342
Db 1261 -----AlaAspIleValLysAsnIleIleGlyAlaAlaGlyGluIle 1274
QY 3343 TATGCCCATGGAAGATATGATGATATCGATGATCGAGGCTATCTCTCGAGGCTATGTTTTCAGTGCAGGA 3402
Db 1275 ValGlyAlaGly-----AspAlaValGln-----GlyIleSerGluGly 1287
QY 3403 AGTAGAGTC 3411
Db 1288 SerAsnIle 1290
```

## RESULT 38

```
US-10-467-657-1088
; Sequence 1088, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: PONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1088
; LENGTH: 1468
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1088
```

Alignment Scores:





```
Db 304 SerHisIleSerSerAlaGlnGlyGlyLeuValThrSerSerSerAspAlaThrIle 323
Qy 1192 AGTTTTCCAGAAATACGCGGTAGAGTTTGATGGGAACGTAGCCGAGTAGGAGGAGG 1251
Db 324 AsnPheSerGlyThrAlaAlaGlnArg-----AsnSer 334
Qy 1252 ATTTACTCTACGGGAACGTTCTTCTCTGAATAATGGAAGAACCTGTTCTCTCAACAAT 1311
Db 335 IlePheSerGlySerThrGlyAlaSerAlaGlnThrAlaThrAlaValIleAsn--- 353
Qy 1312 GTTGCTTCTCTGTTTACATTGCTGCTAGCAACCAACCAAGTGCAGAGCTTCTAATACG 1371
Db 354 -----MetGlnAsnThr 357
Qy 1372 AGTAATAATTACGGAGATCGAGAGCTATCTTCTGTGAAGATGGTGGCAACGAGATCC 1431
Db 358 AspIleThrValAspArgAsnGlySerLeuAlaLeuGlyLeuTrpAlaLeuSerGlyGly 377
Qy 1432 AATAACTCTGGA---TCAGTTTCTTTGATGGAGGAGGTAGTTTCTTTAGTAGCAAT 1488
Db 378 ArgIleThrGlyAspSerLeuAlaIleThrGly----- 388
Qy 1489 GTAGCTGCTGGGAAGGGGAGCTATTATGCC----- 1521
Db 389 ---AlaAlaGlyAlaArgGly---IleTyrAlaMetThrAsnSerGlnIleAspLeuThr 406
Qy 1522 AAAAGCTCTCGGTCTAAGTGTGGCCCTGTACAAATTTTAAAGGAATATCGTAATGAT 1581
Db 407 SerAspLeuValIleAspMetSerThrProAspGlnMetAlaIleAlaThrGlnHisAsp 426
Qy 1582 GTTGGG---GCGATTATTAGGAGATCTCGGAGAGCTCAGTTTATCTGCTGATAT 1635
Db 427 AspGlyTyrAlaAlaSerArgIleAsnAlaSerGlyArgMetLeuIleAsn----- 443
Qy 1636 CGAGATATTATTTCATGGGAATCTTAAAGA----- 1668
Db 444 GlySerValLeuSerIleGlyGlyLeuIleAsnLeuAspMetHisProGlySerValTrp 463
Qy 1669 ACAGCCAAAGAGAAATCTGCGGATGTTAATGGC-----GTAAGTGTGCTCTCA 1716
Db 464 ThrGlySerSerLeuSerAspAsnValAsnGlyGlyLysLeuAspValAlaMetAsnAsn 483
Qy 1717 CAAGCATTTGATGGATCGGAGGAGGAAATAACGACATTAGA----- 1761
Db 484 SerValTrpAsnValThrSerAsnSerAsnLeuAspThrLeuAlaLeuSerHisSerThr 503
Qy 1762 -----GCTAAAGCAGGCGCATCAGATTCTCTTAATGATCCCATC 1800
Db 504 ValAspPheAlaSerHisGlySerThrAlaGlyThrPheThrThrLeuAsn-----Val 521
Qy 1801 GAGATGGCAACCGAAATTAACAGCCAGCGCAGCTTCCAAACTTCTAAATAATACGAT 1860
Db 522 GluAsnLeuSerGlyAsnSerThrPheIleMetArgAlaAspValValGlyGluGlyAsn 541
Qy 1861 GGTGAAGGATACACAGGGATATGTTTGTCTAATGAGAGCAGTACTTTGTACCAAAAT 1920
Db 542 GlyValAsnAsnArgGlyAspLeuLeuAsnIleSerGlySerAlaGlyAsnHisVal 561
Qy 1921 GTTACGATACAGCAA-----GGAAGGATGTTCTTCGT----- 1953
Db 562 LeuAlaIleArgAsnGlnGlySerGluAlaThrThrGlyAsnGluValLeuThrValVal 581
Qy 1954 -----GAAAAGGCAAAATTAATCAGTGAATTTCTTAAGTCAACAGAGTGGAGT 2001
Db 582 LysThrThrAspGlyAlaAlaSerPheSerAlaSerSerGlnValGluLeuGlyGlyTyr 601
Qy 2002 CTGTAT-----ATGGAAGCTGGAGTACATGGATTTGTA----- 2037
Db 602 LeuTyrAspValArgLysAsnGlyThrAsnTrpGluLeuTyrAlaSerGlyThrValPro 621
Qy 2038 -----ACTCCACAACCA-----CCACAACAGCTCTCGCGCTAATCAG 2076
Db 622 GluProThrProAsnProGluProThrProAlaProAlaGlnProPro----- 637
```

---

```
Qy 2077 TTGATCAGCTTTTCCAAATCTGCATTTGTCTCTTCTTCTTTGTAGCAAAACAATGCAGTT 2136
Db 638 -----1le 638
Qy 2137 ACGAATCT---CCTPACCAATCTCCAGCGCAAGATTCATCTCTGCGATCATTTGGTAGC 2193
Db 639 ValAsnProAspProThrProGluProAlaProThrProLysPro-----Thr 654
Qy 2194 ACAACTGCT---GGTCTCTTACAAATTAGTGGGCTTATCTTT----- 2232
Db 655 ThrThrAlaAspAlaGlyGlyAsnTyrLeuAsnValGlyTyrLeuLeuAsnTyrValGlu 674
Qy 2233 -----TTTCAGGATTTGGATGATACAGCTTATGATAGTATGAT 2271
Db 675 AsnArgThrLeuMetGlnArgMetGlyAspLeuArgAsnGlnSerLysAspGlyAsnIle 694
Qy 2272 TGGCTA----- 2277
Db 695 TrpLeuArgSerTyrGlyGlySerLeuAspSerPheAlaSerGlyLysLeuSerGlyPhe 714
Qy 2278 -----GGTTCTAATCAAAAAAATCAATGTCTCGAAA 2307
Db 715 AspMetGlyTyrSerGlyIleGlnPheGlyAspLysArgLeuSerAspValMetPro 734
Qy 2308 TTACAGTTAGGAGCTAAG---CCCCAGCTAATGCCCATCAGATTGTGACTCTAGGGAAT 2364
Db 735 LeuTyrValGlyLeuTyrIleAspSerThrHisAlaSerProAspTyrSerGlyGlyAsp 754
Qy 2365 GAGATGCTCTAAGTATGGCTATCAAGAACTCGAAGCTGCGTGGGATCCTAATACAGCA 2424
Db 755 GlyThrAlaArgSerAspTyrMetGlyMetTyrAla-----SerTyrMetAla 770
Qy 2425 AATAATGGTCTTATCT---CTGAAGCTACATGAGCTAAAGTGGTATAAT 2475
Db 771 GlnAsnGlyPheTyrSerAspLeuValIleLysAlaSerArgGlnLysAsnSerPheHis 790
Qy 2476 CTGGGCTGAGCGAGTAGTCTTTGTTGTTCCAAATAGTTTATGGGATCCATTTTAGAT 2535
Db 791 -----ValLeuAsp 793
Qy 2536 ATACGATCTCGCATTCAGCAATTCAGCAAGTGTGAGTGGCGCTCTTATTGT----- 2589
Db 794 SerGlnAsnAsnGlyValAsnAlaAsnGlyThrAlaAsnGlyMetSerIleSerLeuGlu 813
Qy 2590 -----CGAGGATTATGGTCTTCTGGAGTTTCG 2616
Db 814 AlaGlyGlnArgPheAsnLeuSerProThrGlyTyrGlyPheTyrIleGluProGlnThr 833
Qy 2617 AATTTCTTCTATCATGACCGCATGCTTTAGGTCAGGATATCGGTATATTAGTGGGGT 2676
Db 834 GlnLeuThrTyrSerHisGlnAsnGluMetAlaMet-----LysAlaSerAsnGly 850
Qy 2677 TATTCCTTAGGAGCAAACTCTACTTTGGATCATCGATGTTGGTCTAGCATTTACCGAA 2736
Db 851 LeuAsnIleHisLeuAsnHisTyr-----GluSerLeuGlyArgAla---SerMet 867
Qy 2737 GTATTTGTAGATCTAAAGATATGATGTGTGTCGTTCATCATCATCTGCTGCATAGGA 2796
Db 868 IleLeuGly-----TyrAspIleThrAlaGlyAsnSerGlnLeu----- 880
Qy 2797 TCCGTTTATCTATCTACCCAA-----CAAGCTTTATGTGA-----TCTATTGTTTC 2844
Db 881 AsnValTyrValLysThrGlyAlaIleArgGluPheSerGlyAspThrGluTyrLeuLeu 900
Qy 2845 GAGATGCTGTTATCCGTGCTAGCTACGGGTTT---GGGAAT----- 2883
Db 901 AsnAspSer-----ArgGluLysTyrSerPheLysGlyAsnGlyTyrAsnAsnGlyVal 918
Qy 2884 -----CAGCATATGAAAACCTCATATACATTTCAGAGGAGCGATGTCGT 2931
Db 919 GlyValSerAlaGlnTyrAsnLys---GlnHisThrPheTyrLeuGluAlaAspTyrThr 937
```







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: May 13, 2006, 10:00:08 ; Search time 28 Seconds  
(without alignments)  
4572.014 Million cell updates/sec

Title: US-10-701-844-1  
Perfect score: 7883  
Sequence: 1 999caaaactctcccccg.....gcattcctaagaagaattc 4435

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DRV=xlp  
-O=/abes/ABSSWEB spool/US10701844/runat 12052006 165423 26326/app query.fasta\_1  
-DB=PIR -QFMT=faetan -SUFFIX=trp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02p  
-USER=US10701844 @CNG 1.1.63 @runat 12052006 165423 26326 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5311.5	65.1	1013	2	G71460
2	3846	48.8	987	2	H81722
3	1589	20.2	1016	2	H71460
4	1366	17.3	995	2	C81593
5	1365.5	17.3	973	2	H86547
6	1365.5	17.3	973	2	F72076
7	1179	15.0	983	2	H81723
8	1101	14.0	928	2	H86546
9	1101	14.0	928	2	H81591
10	1092	13.9	949	2	F81591
11	1090	13.8	928	2	D72077
12	1090	13.8	928	2	H86546
13	1083	13.7	930	2	D86546
14	1083	13.7	930	2	H81591

15	1079	13.7	930	2	D72078	polymorphic outer
16	1041	13.2	936	2	C72078	polymorphic outer
17	1040	13.2	936	2	B81591	polymorphic membra
18	1040	13.2	936	2	C86546	polymorphic outer
19	1021	13.0	928	2	B72077	polymorphic membra
20	1021	13.0	928	2	E86546	polymorphic outer
21	1015	12.9	1276	2	B86546	polymorphic outer
22	1015	12.9	1276	2	C81591	polymorphic membra
23	991.5	12.6	1407	2	B72078	polymorphic outer
24	979.5	12.4	772	2	H86492	Pmp 3 [imported] -
25	973	12.3	922	2	B72131	polymorphic outer
26	973	12.3	922	2	E86491	polymorphic outer
27	972	12.3	922	2	F81539	polymorphic membra
28	939.5	11.9	841	2	B72130	polymorphic membra
29	789	10.0	867	2	F81721	polymorphic membra
30	771.5	9.8	878	2	B71460	probable outer mem
31	703	8.9	712	2	E86492	polymorphic outer
32	702	8.9	978	2	G72076	polymorphic outer
33	697	8.8	978	2	C86547	polymorphic membra
34	697	8.8	978	2	B81593	polymorphic membra
35	679.5	8.6	976	2	F81722	polymorphic membra
36	649.5	8.2	964	2	E71460	probable outer mem
37	644	8.2	938	2	F86548	polymorphic outer
38	644	8.2	938	2	H72074	polymorphic membra
39	611	7.8	1520	2	H81731	polymorphic membra
40	609	7.7	946	2	C86549	polymorphic outer
41	609	7.7	946	2	D81594	polymorphic membra
42	607	7.7	946	2	C72075	polymorphic outer
43	599.5	7.6	947	2	G86557	polymorphic membra
44	599.5	7.6	947	2	D72067	polymorphic membra
45	597.5	7.6	1531	2	H71468	probable outer mem

ALIGNMENTS

RESULT 1

G71460

probable outer membrane protein G - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
C:Accession: G71460  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra

A:Reference number: A71570; MUID:99000809; PMID:9784136  
A:Accession: G71460  
A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1013 <ARN>

A:Cross-references: UNIPROT:O84879; UNIPARC:UPI0000131CF5; GB:AE001360; GB:AE001273; NI

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: pmpG

Alignment Scores:

Pred. No.: 0 Length: 1013  
Score: 5311.50 Matches: 983  
Percent Similarity: 98.8% Conservative: 18  
Best Local Similarity: 97.0% Mismatches: 11  
Query Match: 65.1% Indels: 1  
DB: Gaps: 1

US-10-701-844-1 (1-4435) x G71460 (1-1013)

QY 382 ATGCAACGCTCTTTCCATAGTTCTTTCTTCAATGATTCTAGCTATTCTTCTGCTCT 441

Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaIleSerCysCys 20

QY 442 TTAATATGGGGGGGATATGACGACAGAAATCATGTTCTCTCAAGGATTTACGATGGGAG 501

Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40

QY 502 ACGTAACTGTATCATTTCCCTATCTACTATTAGGAGATCCGAGTGGGACTACTGTTTTT 561

Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTGAGAGAGTTAAACATTAATAAATCTTGACAAATCTATTGACAGCTTTGCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
Qy 622 TGTTTGGGAACCTTAATTAGGAGTTTACTGTTTGGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy 682 GAGAACATACGAGCTTCTACAAATGGCGAGCTCTAACTAATAGCGCTGCTGAGGACTG 741  
Db 101 GluAsnIleAlaGThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
Qy 742 TTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCCAAATTCGAAATTCATTACTTGCCTGA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
Qy 802 CTGCTGCTGCAACACTAATAAGGTAGCCAGACTCCGACGACAACTCTACACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
Qy 862 AATGGTACTATTATTCTTAAACAGACTCTTTGTTACTCAATAATGACAGATTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
Qy 982 GGAATTAGCAAGCTTTGTCTTCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys 220  
Qy 1042 CAAAGTACCAAGCTTCTCTGCTAGCTTAAGCTTAAGAGGCTCTATTGCTTGTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
Qy 1102 GTTCAGAGATGAAGGGGGAGGATGCTGCTGTTTCAGGATGGCGAGGAGGTGCA 1161  
Db 241 ValAlaGlyValArgGlyGlyIleAlaValAlaValGlnAspGlyGlnGlyValSer 260  
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTGCGGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
Qy 1222 GATGGGAAGTACCGGAGTAGGAGGGATTTACTCTCTACGGGAAGCTGCTTCTCTG 1281  
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
Qy 1282 AATAATGGAACCTTGTCTCAACAAATGTTGCTTCTCTGTTTACATTTGCTGCTAAG 1341  
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320  
Qy 1342 CAACCAACAAGTGGACAGCTTCTAATACAGTAATAATATTACGAGATGGAGGACTATC 1401  
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
Qy 1402 TTTCTGTAGAATGGTGGCAA---GCAGATCCAAATACTCTGGATCAGTTTCTTTGAT 1458  
Db 341 PheCysLysAsnGlyAlaGlnAlaLysSerAsnAsnSerGlySerValSerPheAsp 360  
Qy 1459 GAGAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518  
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380  
Qy 1519 GCCAAAAGCTCTCGGTGCTAATCTGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAAT 1578  
Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
Qy 1579 GATGCTGGAGCATTTATTAGAGAACTCTGGAGAGCTCAGTTTATCTGCTCATTTATGGA 1638

Db 401 AspGlyGlyAlaIleTyrLeuGlyLysSerGlyLeuLeuSerLeuSerAlaAspTyrGly 420  
Qy 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAAATGCTGCCGATGTTAAT 1698  
Db 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
Qy 1699 GGCGTAATCTGTCCTCACAGCCATTTCCGATCGGATCGGGAGGAAATAACACACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
Qy 1759 AGACTAAAGCAGGGCATCAGATTCTCTTAATATCCCATCGAGATGGCAACGGAAT 1818  
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
Qy 1819 AACGACGACGCGAGCTTCTCCAACTTCTAAAAATTAACGATGTGAAGGATACACAGG 1878  
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyGlyTyrThrGly 500  
Qy 1879 GATATTCTTTTTCGTAATGGAGCAGTACTTTGACCAAAATGTTACGATAGACCAAGGA 1938  
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
Qy 1939 AGGATTGTTCTCTGAAAAGGCCAAATTAATCAGTGAATCTCTAAGTCAGACAGGTGGG 1998  
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
Qy 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCAACACAG 2058  
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
Qy 2059 CCTCTCGCGCTATCAGTTGATCAGCTTCTCCAACTGTCGATTTGCTCTTCTCTTTG 2118  
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
Qy 2119 TTAGCAACAATGCAGTTTACGAATCTCTACCAATCTCCAGCGCGAAGATTCATCCT 2178  
Db 581 LeuAlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600  
Qy 2179 GCAGTCAATGTTAGCACAACTGCTGTTTCTGTTTCAATTAGTGGGCTTATCTTTTQAG 2238  
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
Qy 2239 GATTTGGATGATACAGCTTATGATAGTATGATGCTGCTAGGTCTTAATCAAAAATCAAT 2298  
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640  
Qy 2299 GTCTGAAAATACAGTTAGGAGTAAAGCCCGCAGCTAATGCCCATCAGATTTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
Qy 2359 GGGATAGAGTGCCTAAGTATGGCTATCAAGGAAGCTGGAAAGCTTGGGATCTTAAT 2418  
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680  
Qy 2419 ACAGCAATAATGCTCTTACTACTGAAAGCTACATGAGCTAAACCTGGGTATAATCCT 2478  
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro 700  
Qy 2479 GGGCTGAGCAGTAGCTTCTTTGTTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720  
Qy 2539 CGATCTGCGCATTCAGCAATTAAGCAAGTGTGGAAGTGGCGCTCTTATTTGCGAGGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
Qy 2599 TGGGTTTCTGAGTTTCGAATTTCTTCATCATGACCGGATGCTTTTAGGTCAGGATAT 2658  
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
Qy 2659 CGGTATATTAGTGGGGTTTATCTCTTAGAGCAAACTCTCTATTTTGGATCATCGATGTT 2718  
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780

QY 2719 GGTCTAGCATTTTACCGAAGTATTGGTAGATCTAAGATTATGCTAGTGCTGTTCCAAAT 2778  
 Db GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
 QY 2779 CATCATGCTTGGCATAGGATCCGGTTTATCTATCTATCCCAACAAAGCTTTATGTGATCCTAT 2838  
 Db HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
 QY 2839 TTGTTCCGAGATCGCTTTATCCGTGTAGCTACGCTCGGTGGGATCGGATCAGCATATGAACCC 2898  
 Db LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
 QY 2899 TCATATACATTTCCAGAGAGACGATGTTGCTGGGTAATAACTGTCTGGCTGGAGAG 2958  
 Db SerTyrThrPheAlaGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860  
 QY 2959 ATTGGAGCGGGATTACCGATTGTGATTCTTCACTTAAGCTCTATTGAATGAGTTGCGT 3018  
 Db IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
 QY 3019 CCTTTCGTCGAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGGAT 3078  
 Db ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
 QY 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCATCTCTGTTGGAGTGAAG 3138  
 Db GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
 QY 3139 TTTGATCGATGTTCTTAGTACACATCTCTAATAAATATAGCTTTTATGCGCGCTATATCTGT 3198  
 Db PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
 QY 3199 GATGCTTATGCAACCATCTCTGTACTGAGACAAACGCTCTCATCCCATCAAGAGACATGG 3258  
 Db AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
 QY 3259 ACAACAGATCCCTTTCATTAGCAACATGGAAGTGTGCTTGTAGAGATCTAGTATGCT 3318  
 Db ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
 QY 3319 TCTCTAACAGTAAATAGAGATATATGCGCATGGAGATATGATATCCAGATGCTTCT 3378  
 Db SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
 QY 3379 CGAGGCTATGTTGATGTCAGGAAGTAGAGTCCGGTTC 3417  
 Db ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013  
 RESULT 2  
 H81722  
 polypeptide membrane protein G family TC0263 [imported] - Chlamydia muridarum (strain N1  
 C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C;Accession: H81722  
 R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
 A;Reference number: A81500; MUID:20150255; PMID:10684935  
 A;Accession: H81722  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-987 <TET>  
 A;Cross-references: UNIPROT:Q9PL45; UNIPARC:UPI0000057867; GB:AE002293; GB:AE002160; NII  
 A;Experimental source: strain Nigg (MoPn)  
 C;Genetics:  
 A;Gene: TC0263  
 Alignment Scores:  
 Pred. No.: 8,25e-257 Length: 987  
 Score: 3845.00 Matches: 739  
 Percent Similarity: 82.9% Conservative: 106

Best Local Similarity: 72.5% Mismatches: 136  
 Query Match: 48.8% Indels: 38  
 DB: 2 Gaps: 13  
 US-10-701-844-1 (1-4435) x H81722 (1-987)  
 QY 379 GTGATGCAAAACGCTCTTTCCATAGTCTTTTCTTTCAATGATTCTAGCTTATCTTCTGCTGC 438  
 Db 1 MetMetGlnThrProPheHisLysPhePheLeuLeuAlaMetLeuSerTyr ----- 17  
 QY 439 TCTTTAAATGGGGGATATGCACAGAAATCATGGTCTCTCAAGGAATTTACGATGGG 498  
 Db 18 SerLeuLeuGlnGlyHisAlaAlaAspIleSerMetProProGlyIleTyrAspGly 37  
 QY 499 GAGACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTT 558  
 Db 38 ThrThrLeuThrAlaProPheProTyrThrValIleGlyAspProArgGlyThrLysVal 57  
 QY 559 TTTTCTGCAAGAGATTAAACATAAATACTTGCAATCTTATTCGACGTTTGCCTTTA 618  
 Db 58 ThrSerSerGlySerLeuGluLysAsnLeuAspAsnSerIleAlaThrLeuProLeu 77  
 QY 619 AGTGTGTTGGGAACCTTATAGGAGTCTTCTGTTTATAGGAGAGACACTCGTGTACT 678  
 Db 78 SerCysPheGlyAsnLeuGlyAsnPheThrIleAlaGlyArgGlyHisSerLeuVal 97  
 QY 679 TTCGAGAACATACGAGCTTCTCAAAATGGGAGCTCTAAAGTAATAGCGCTGCTGATGGA 738  
 Db 98 PheGluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnHisAlaProSerGly 117  
 QY 739 CTGTTTACTTATGAGGGTTTAAAGAAATTTATCCTTTTCCAATTGCAATTCATTACTTGGC 798  
 Db 118 LeuPheValIleGluAlaPheAspGluLeuSerLeuLeuAsnCysAsnSerLeuValSer 137  
 QY 799 GTATCGCTGCTGCAACGACTTAATAGGGTAGCCAGACT---CCGACGCAACATCTACA 855  
 Db 138 ValValPro-----GlnThrGlyGlyThrThrThrSerVal 149  
 QY 856 CCGTCTAATGGTACTATTATTTCTAAACACAGACTTTTGTGTACTCAATATAGAACTTC 915  
 Db 150 ProSerAsnGlyThrIleTyrSerArgThrAspLeuValLeuArgPheIleLysVal 169  
 QY 916 TCATTCTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGACTTAAACG 975  
 Db 170 SerPheTyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaGlnSerLeuMet 189  
 QY 976 GTTCAAGGAATTAGCAAGCTTTGTCTTCCAAAGAAAATACTGCTCAAGCTGATGGGGA 1035  
 Db 190 ValAsnGlyIleGluLysLeuCysThrPheGlnGluAsnValAlaGlnSerAspGlyGly 209  
 QY 1036 GCTTGTCAAGTAGTACCAGTTTCTCTGCTATGCTTAAACGAGGCTCCTATTGCTTTGTA 1095  
 Db 210 AlaCysGlnValThrLysThrPheSerAlaValGlyAsnLysValProLeuSerPheLeu 229  
 QY 1096 GCCAATGTTGACGAGTAGAGGGGAGGATGCTGCTTTCAGGATGGGACGAGGGA 1155  
 Db 230 GlyAsnValAlaGlyAsnLysGlyGlyValAlaAlaValLysAspGly---GlnGly 248  
 QY 1156 GTGTCATCATCTACTTCAACAGAGATCCAGTAGTAGTGTGTTTCCAGAAATATCTCGGTA 1215  
 Db 249 AlaGlyGlyAlaThr-----AspLeuSerValAsnPheAlaAsnAsnThrAlaVal 265  
 QY 1216 GAGTTTCATGGGAACGTAGCCGAGTAGGAGGAGGATTTTACTCTACGCGGAGCTGTGCT 1275  
 Db 266 GluPheGluGlyAsnSerAlaArgIleGlyGlyIleTyrSerAspGlyAsnIleSer 285  
 QY 1276 TTCCTGAATAATGAAAAACCTTGTCTCAACAAATGTTCTCTCTCTCTCTCTCTCTCTCT 1332  
 Db 286 PheLeuGlyAsnAlaLysThrValPheLeuSerAsnValAlaSerProIleTyrValAsp 305  
 QY 1333 ---GCTGCTAAGCAACCAACAGTGGACAGGCTTCTTAATACGAGTAATAATTACGGAGAT 1389  
 Db 306 ProAlaAla-----AlaGlyGlyGlnProProAlaAspLysAspAsnTyrGlyAsp 322

```
QY 1390 GGAGGAGCTATCTCTGTGAAGATGGTGGCAAGCAGGATCCAATAACTCTGGATCAAGTT 1449
Db GlyGlyAlaIlePheCysLysAsn-----AspThrAsnIleGlyGluVal 337
QY 1450 TCCTTTGATGAGAGGAGTATGTTCTTTAGTAGCAATGATGCTGGGAAAGGGGGA 1509
Db SerPheLysAspGluGlyValValPhePheSerLysAsnIleAlaAlaGlyLysGlyGly 357
QY 1510 GCTATTTATGCCAAAAGCTCTCGGTTGCTTAACGTGTCGCCCTGTACAAATTTTAAAGGAAT 1569
Db AlaIleTyrAlaLysLysLeuThrIleSerAspCysGlyProValGlnPheLeuGlyAsn 377
QY 1570 ATCGCTAATGATGGGAGCGCATTTATTTAGAGAAATCTGAGAGCTCAGTTTATCTGCT 1629
Db ValAlaAsnAspGlyAlaIleTyrLeuValAspGlnGlyLeuSerLeuSerAla 397
QY 1630 GATTATGGAGATATTTATTCGATGGGAATCTTAAAGNACAGCCAAAGAGAAATGCTGCC 1689
Db AspArgGlyAspIlePheAspGlyAsnLeuLysArgMetAlaThrGlnGlyAlaAla 417
QY 1690 GATGTTAATGGCTAACTGTGCTCCACAAAGCCATTTCCGATGGGATCGGAGGAAATA 1749
Db ThrValHisAspValMetValAlaSerAsnAlaIleSerMetAlaThrGlyGlyGlnIle 437
QY 1750 ACGACATTAAAGAGCTAAAGCGGCATCAGATTCCTTTAATGATCCCATCGAGATGGCA 1809
Db ThrThrLeuArgAlaLysGluGlyArgArgIleLeuPheAsnAspProIleGluMetAla 457
QY 1810 AACGAAATACACGCGCAGCTCTCCAAACTTCTAAAAATTAACGATGGTGAAGGA 1869
Db AsnGly-----GlnProVal-----IleGlnThrLeuThrValAsnGluGlyGly 473
QY 1870 TACACAGGGATATGTTTCTTAATGAAGCAGTACTTTGTACAAAATGTTACGATA 1929
Db TyrThrGlyAspIleValPheAlaLysGlyAspAsnValLeuTyrSerSerIleGluLeu 493
QY 1930 GAGCAGGAGGAGTGTCTTCGTGAAAGGCAAAATATACAGTGAATTCCTAAAGTCAG 1989
Db SerGlnGlyArgIleIleLeuArgGluGlnThrLysLeuValAsnSerLeuThrGln 513
QY 1990 ACAGTGGGAGTCTGATATGAAGCTGGGAGTACATGGGATTTGTAACTCCACACCA 2049
Db ThrGlyGlySerValHisMetGluGlyGlySerThrLeuAspPheAlaVal----- 530
QY 2050 CCACAACAGCTCTCGCGCTAACTAGTTGATCAGCTTCCAAATCTGCATTTGTCTCT 2109
Db ---ThrThrProProAlaAlaAsnSerMet---AlaLeuThrAsnValHisPheSerLeu 548
QY 2110 TCTCTTTGTTAGCAAAATGCGATTACGAATCTCTCCATCCATCTCCAGCGCAAGAT 2169
Db AlaSerLeuLeuLysAsnAsnGlyValThrAsnProProThrAsnProProValGlnVal 568
QY 2170 TCTCATCTCGAGTCAATGTTAGCACAACTGCTGTTCTGTTACAAATAGTGGGCTATC 2229
Db SerSerProAlaValIleGlyAsnThrAlaAlaGlyThrValThrIleSerGlyProIle 588
QY 2230 TTTTGTGAGATTGGATCATACAGCTTATGATAGTATGATGATGTTGCTTCTAATCAA 2289
Db PhePheGluAspLeuAspGluThrAlaTyrAspAsnAsnGlnThrPheGlyAlaAspGln 608
QY 2290 AAAATCAATGCTCTGAAATTTACAGTTAGGAGCTTAAGCCCCCAGCTAATCCCCATCAGAT 2349
Db ThrIleAspValLeuGlnLeuHisLeuGlyAlaAsnProProAlaAsnAlaProThrAsp 628
QY 2350 TTGACTCTAGGGAATGAGATGCCATAGTATGGTATCAAGGAGCTGGAGCTTGGTGG 2409
Db LeuThrLeuGlyAsnGluSerSerLysTyrGlyTyrGlnGlySerThrThrLeuGlnThr 648
QY 2410 GATCCTTAATACAGCAAT-----AATGGTCTTATATCTCTGAAAGCTACATCGACT 2460
Db GluProAspProAlaAsnProProGlnAsnAsnSerTyrMetLeuLysAlaSerThrThr 668
```

```
QY 2461 AAAACTGGGTATAATCCTGGGCGCTGACGAGTAGCTTCTTTGGTTCCAAATAGTTTATGG 2520
Db LysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValSerAsnSerLeuTrp 688
QY 2521 GGATCCCAATTTAGATATACGATCTGCGCATTTCAACAATTCGAAGTGTGTGATGGGCGC 2580
Db GlySerIleLeuAspValArgSerAlaHisSerAlaIleGlnAlaSerIleAspGlyArg 708
QY 2581 TCTTATTGTCAGGATATATGGGTTTCTGGAGTTTCGAATTTCTTCTATCATGACCGCAT 2640
Db AlaTyrCysArgGlyIleTrpIleSerGlyIleSerAsnPhePheTyrHisAspGlnAsp 728
QY 2641 GCTTTAGTCAAGGATATCGGTATATAGTGGGGTTATTCTCTTAGGAGCAAACTCCATC 2700
Db AlaLeuGlyGlnGlyTyrArgHisIleSerGlyGlyTyrSerIleGlyAlaAsnSerTyr 748
QY 2701 TTTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTATTTGGTAGATCTAAAGATTAT 2760
Db PheGlySerSerMetPheGlyLeuAlaPheThrGluThrPheGlyArgSerLysAspTyr 768
QY 2761 GTAGTGTGCTGCTTCCAAATCATCATGCTTGCATAGATCCGTTTATCTATCTACCCAA 2820
Db ValValCysArgSerAsnAspHisThrCysValGlySerValTyrLeuSerThrArgGln 788
QY 2821 GCTTTATGTCGATCCTATTGTTTCGAGATCGGTTTATCCGTGCTAGCTACGGGTTTGG 2880
Db AlaLeuCysGlySerCysLeuPheGlyAspAlaPheValArgAlaSerTyrGlyPheGly 808
QY 2881 AATCAGCATATGAAAACTCATATACATTTTCAGAGGAGAGCGATGTTCTGTTGGGATAAT 2940
Db AsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAsnValArgTyrAspAsn 828
QY 2941 AACTGTCTGCTGAGAGATTTGGAGCGGATTCAGATTGTGATCTCCATCTCAAGCTC 3000
Db AsnCysValValGlyGluValGlyAlaGlyLeuProIleMetLeuAlaAlaSerLysLeu 848
QY 3001 TATTTGAATGATGTCGCTTCCTTCGCAAGCTCAGTTTCTTATGCGCATCATGAATCT 3060
Db TyrLeuAsnGluLeuArgProPheValGlnAlaGluPheAlaTyrAlaGluHisGluSer 868
QY 3061 TTTACAGAGGAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTAAATCTATCA 3120
Db PheThrGluArgGlyAspGlnAlaArgGluPheLysSerGlyHisLeuMetAsnLeuSer 888
QY 3121 GTTCTGTGTGAGTGAAGTTGATGATGTTCTAGTACACATCTCTTAATAATATAGCTTT 3180
Db IleProValGlyValLysPheAspArgCysSerSerLysHisProAsnLysTyrSerPhe 908
QY 3181 ATGGCGGCTTATCTGATGCTTATCGCACCATCTCTGTTACTGAGACAGCTCCTA 3240
Db MetGlyAlaTyrIleCysAspAlaTyrArgSerIleSerGlyThrGluThrThrLeuLeu 928
QY 3241 TCCCATCAAGAGACATGAGCAACAGATCGCTTTTCAATTTAGCAAGACATGGAGTTTGGTT 3300
Db SerHisLysGluThrThrThrAspAlaPheHisLeuAlaArgHisGlyValMetVal 948
QY 3301 AGAGATCTATGATGCTTCTCTCAACAGTAAATATAGAAGTATATAGGCCATGGAAGATAT 3360
Db ArgGlySerMetTyrAlaSerLeuThrGlyAsnIleGluValTyrGlyHisGlyLysTyr 968
QY 3361 GAGTATCGAGATGCTTCTCGAGGCTATGTTTGTAGTCGAGGAGTAGAGTCGCGTTC 3417
Db GluTyrArgAspAlaSerArgGlyTyrGlyLeuSerIleGlySerLysIleArgPhe 987
```

## RESULT 3

H71460  
probable outer membrane protein H - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
R:Accession: H71460  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis



A;Reference number: A71570; MUID:99000809; PMID:9784136  
A;Accession: H71460  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1016 <ARN>  
A;Cross-references: UNIPROT:O84880; UNIPARC:UPI0000046AC7; GB:AE001360; GB:AE001273; NID  
A;Experimental source: serotype D, strain UW-3/Cx  
C;Genetics:  
A;Gene: pmph  
C;Superfamily: Polymorphic membrane protein H family

Alignment Scores:		
Pred. No.:	2.3e-101	Length: 1016
Score:	1589.00	Matches: 311
Percent Similarity:	92.7%	Conservative: 5
Best Local Similarity:	91.2%	Mismatches: 10
Query Match:	20.2%	Indels: 15
Gap:	2	Gaps: 2

US-10-701-844-1 (1-4435) x H71460 (1-1016)

Qy	3451	ATGCCTTTTCTTTGAGATCTACATCATTTTGTGTTTTTGTGTTTGTGTTCTATTTCG	3510
Db	1	MetProPheSerLeuAlaSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrSer	20
Qy	3511	TATCGATTCCGAGACTCTCCTCAAGTGTTAACGGCTAATGTTAACCACTCTCTTTTAAGGGA	3570
Db	21	TyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrProPheLeuGly	40
Qy	3571	GACCATGTTTACTTGATGGAGACTCGCTTTTCTCAATGCTCTATCGAGAGCTGAAGAA	3630
Db	41	AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn	60
Qy	3631	GGTTCGATTATCTCAGCTTAATGCGGCACAAATTAACAGATTACCGACACAAACCATACATTA	3690
Db	61	GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu	80

QY	3751	ACACTTACTCTCGAGAGATTTTTCAGAGTCGTGATGTTCTCGAAAAAATGTTTCTTCCGGAGAA	3810
DB	101	ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu	120
QY	3811	AAGGGAATGATCTCCGGGAAAACCGTGTAGTATTTCCGGAGCAGCGCAAGTCATTTCTCG	3870
DB	121	LysGlyMetIleSerGlyLeuThrValSerIleSerGlyAlaGlyGluValIlePheTrp	140

[illegible]

Qy	3955	AAAGGGTCTATATTTTCTGTACAGACTAGTTGGAGATCTCAGGCGTCACAAAAAGGGGTC	4011
Db	181	LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLysGlyVal	200
Qy	4015	ATGTTTCGATAATAATGTCGGGAATTTCCGACAGATTTTTTCAGAGGTAAAGATAATAAT	4074
Db	201	MetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsn	220

QY	4135	CTGTAAAGGGAAGTTCTTTTTCACAGATAACGTAGCCTCTTCGGAGCGCGGAGTGTTTA	4194
Db	238	nCybLy8GlyLy8ValSerPheThrAspAsnValAlaSerCy8GlyGlyGlyValVal	258
QY	4075	GCTGGTGGTGGAGGACAGTGGTTCCTGCATACCATCAAGATACGACCTTTTACAGTTATATA	4193
Db	221	AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLy8As	238

Qy	4195	TAAAGCATGTCTCTTTCAAAGACAATGAAGAGGCATATCTTCGAGGGAACACAGC	4253
Db	258	rLyGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAl	278
Qy	4255	ATACGATGATTAAAGGATCTCTGCTGCTACTAATCAGGATCAGAAATACGGAGACAGAGG	4314
Db	278	aTyAspAspLeuGlyIleLeuAlaIleThrSerArgAspGlnAsnThrGluThrGlyGly	298
Qy	4315	CGGTGAGGAGTATTGTCCTCCAGATGATCTGTAAAGCTTTGMAAGGCATAAAGCTTC	4374
Db	298	yGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlyLys	318
Qy	4375	TATTGTTTTTGATTACAACTTTGCAAAGGCAGAGCGCGGAAGCATCTCTAACCAAGAATT	4434
Db	318	rIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPh	338
Qy	4435	C 4435	
Db	338	e 338	

RESULT 4

C81593  
polymorphic membrane protein G family CP0299 [imported] - Chlamydophila pneumoniae  
C/Species: Chlamydophila pneumoniae, Chlamydia pneumoniae  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 15-Jun-2001  
C/Accession: C81593  
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gall, S.R.; Heidelberg, J.F.; White, O.;  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.;  
Nucleic Acids Res. 28, 1397-1406, 2000  
A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae  
A/Reference number: 8A1500. MIMD:20150255: PMID:10684935

A:Accession: C81593  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-595 <REA>  
A:Cross-references: UNIPARC:UPI00001655F9; GB:AE002161; NID:G71899  
A:Experimental source: strain AK39, HL cells  
A:GeneticIS:  
A:Gene: C81593  
A:Gene: C81593  
*Chlamydomonas reinhardtii* outer membrane protein G

Alignment Scores:	
Pred. NO.:	5.23e-86
Score:	1366.00
Percent Similarity:	49.4%
Best Local Similarity:	33.3%
Query Match:	17.3%
DB:	2
Alignments:	29
Matches:	356
Conservative:	172
Mismatches:	425
Indels:	116
Gaps:	29

UUS-10-707-844-1 (1-4435) x C81593 (1-995)

QY	331	TTAAATTTTATGAAGCGGAGTAATAATTTATCTCTC-----	369
ph			
		LeuPhePheTyrSer-----LeuSerLeuLeuCysArgileileIerophe 17	

QY 370 ---AGCTTTTGTGTGATGCACACGCTCTCTTCCCAATAGTTCCTTTCTTTCAATGATTCAGCT 426

427 TATTCTTGTCTCTTTAAATGGGGGGGATATGCAGCAGAAATCATGTTCCTCAAGGA 486

487 ATTTACGATGGGGAGACGTTAACTGTATCATTTCCCTATACTGTGTATAGGAGATCCGAGT 546

QY 547 GGGACTACTGTTTTTCTGCAGGAGAGTTAAACATTTAAAAATCTTTGACAAATTTCTATTGCA 606

QY 607 GCCTTGCCCTTAAGTTGTTTTGGGAACCTATTAGGAGATTCTCTTTAGGGGAGGGA 666

Db 94 ArgThrSerSerSerCysPheSerAsnArgAlaGlyAlaLeuGluInLeuGlyLysGly 113  
Qy 667 CACTCGTTGACTTCGAGAACATACGACTCTTACAAATGGGGAGCTCTAAGTAATAGC 726  
Db 114 GlyValPheSerPheLeuAsnIleArgSerSerAlaAspGlyAlaIleSerSerVal 133  
Qy 727 GCTGCTGATGCA-----CTGTTTACTATTGAGGGTTTTTAAAGAAATATCTCTTT 774  
Db 134 IleThrGlnAsnProGluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePhe 153  
Qy 775 TCCAAATTGCAATTCATTACTTTGCCGTACTGCTGCTGCAACGACTAAATAAGGGTAGCCAG 834  
Db 154 AsnAsnGluSerLeuThr-----SerAsp 162  
Qy 835 ACTCGAGCAGCAACATCTACACCGTCTAATGGTACTATTATTATCTTAAACAGACTCTTTTG 894  
Db 163 ThrSerAlaSerAsnValIleProHisAlaSerAlaIleTyAlaThrThrProMetLeu 182  
Qy 895 TTACTCAATAATGAGAAGTCTCATTCCTATAGTAAATTTAGTCTCTGGAGATGGGGAGCT 954  
Db 183 PheThrAsnAsnAspSerIleLeuPheGlnTyrAsnArgSerAlaGlyPheGlyAlaAla 202  
Qy 955 ATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTCTCCAGAAAT 1014  
Db 203 IleArgGlyThrSerIleThrIleGluAsnThrLysLysSerLeuLeuPheAsnGlyAsn 222  
Qy 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTCTCTGCTATGGCTTAAC 1074  
Db 223 GlySerIleSerAsnGlyGlyAlaLeuThrGlySerAlaAlaIleAsnLeuIleAsnAsn 242  
Qy 1075 GAGGCTCTATTGCTTTGACGAATCTTCAGGAGTAAGGGAGGAGGAGGAGTCTGCT 1134  
Db 243 SerAlaProValIlePheSerThrAsnAlaThrGlyIleTyrGlyGlyAlaIleTyrLeu 262  
Qy 1135 GTTCAGGATGGCAGCAGGAGTGTCACTCATCTACTTCAACAGAAGATCCAGTAGTAGT 1194  
Db 263 Thr-----GlyGlySerMetLeuThrSer----- 270  
Qy 1195 TTTTCCAGAAATACGCGGTGAGTTTGATGGGAACGTAGCCCGAGTAGGAGGAGGAT 1254  
Db 271 --GlyAsnLeuSerGlyValLeuPheValAsnAsnSerSerArgSerGlyGlyAlaIle 289  
Qy 1255 TACTCTACGGAAGCTTCTCTGAATATGGAATAACCTTGTTCTCAACAATGTT 1314  
Db 290 TyrAlaAsnGlyAsnValThrPheSerAsnAsnSerAspLeuThrPheGlnAsnAsnThr 309  
Qy 1315 GCTTCTCTCTGTT---TACATTGCTCTAAGCAACCAACAGTGGACAGGCTTCTAATACG 1371  
Db 310 AlaSerProGlnAsnSerLeuProAlaProThrProProProProProAlaValThr 329  
Qy 1372 AGTAATAATTACGGAGATGGAGAGCTATCTCTTAAGATGGTGGCGAAGCAGGATCC 1431  
Db 330 Pro----LeuLeuGlyTyrGlyGlyAlaIlePheCysThrProProAlaThrProProPro 348  
Qy 1432 AATAACTCTGGATCAGTTTCTCTTTCATGGAGGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
Db 349 ThrGlyVal---SerLeuThrIleSerGlyGluAsnSerValThrPheLeuGluAsnIle 367  
Qy 1492 GCTGCTGGGAAGGGAGCTATTATGCCAAAAGCTCTGGTTGCTAACTGTGGCCCT 1551  
Db 368 AlaSerGluGlnGlyGlyAlaLeuTyrGlyLysLysIleSerIleAspSerAsnLysSer 387  
Qy 1552 GTACAAATTTTAAAGAAATGCTAATGATGGTGGAGGATTTATTTAGGAGAACTCGGA 1611  
Db 388 ThrIlePheLeuGlyAsnThrAlaGlyLysGlyGlyAlaIleAlaIleProGluSerGly 407  
Qy 1612 GAGCTCAGTTATCTGCTGATATCGAGATATTATTTTCATGGGAATCTTAAAGACA 1671  
Db 408 GluLeuSerLeuSerAlaAsnGlnGlyAspIleLeuPheAsnLysLeuSerIleThr 427  
Qy 1672 GCCAAAGAGAATGCTGCCGATGTTAATGGCGTAATGTGTCTCTCAAGCCATTTCCGATG 1731  
Db 428 SerGly-----ThrProThrArgAsnSerIleHisPhe 438

Qy 1732 GGATCGGAGGGAATAACGACATTAAAGAGCTAAAGCAGGCGCATCAGATTCTCTTTAAT 1791  
Db 439 GlyLysAspAlaLysPheAlaThrLeuGlyAlaThrGlnGlyTyrThrLeuTyrPheTyr 458  
Qy 1792 GATCCCATCAGATGGCAACGGAATAACACGACGCGAGCTTCTCAAACTTCTAAAA 1851  
Db 459 AspProIle-----ThrSerAspAspLeuSerAlaAlaSerAlaAlaThrValVal 476  
Qy 1852 ATTAAC-----GATGGTGAAGGATACACACGCGGATATTTGTTTTT----- 1890  
Db 477 ValAsnProLysAlaSerAlaAspGly---AlaTyrSerGlyThrIleValPheSerGly 495  
1891 -----GCTAATCGAAGCAGTACTTTGTACCAA 1917  
Db 496 GluThrLeuThrAlaThrGluAlaAlaThrProAlaAsnAlaThrSerThrLeuAsnGln 515  
Qy 1918 AATGTTACGATAGACGAGGAAGGATTTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAAT 1977  
Db 516 LysLeuGluLeuGluGlyGlyThrLeuAlaLeuArgAsnGlyAlaThrLeuAsnValHis 535  
Qy 1978 TCTCTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGGAGTACATGGATTTT 2034  
Db 536 AsnPheThrGlnAspGluLysSerValIleMetAspAlaGlyThrThr-----Leu 553  
Qy 2035 GTAACTCCACCAACCCACCAAGCCTCTCCGCTAATCAGTTGTATGATCAGCTTTTCCAAT 2094  
Db 554 AlaThrThrAsnGlyAlaAsnAsnThrAspGlyAla-----IleThrLeuAsnLys 570  
Qy 2095 CTGCATTGTCTCTCTCTCTCTTTCTTTAGCAACCAATGCAGTTACGAATCTCTTACCAAT 2154  
Db 571 LeuValIleAsnLeuAspSerLeuAspGlyThrLysAla----- 583  
Qy 2155 CTCCACGCGCAAGATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2211  
Db 584 -----AlaValValAsnValGlnSerThrAsnGlyAlaLeu 595  
Qy 2212 ACAATTAGTGGCTATCTTTTGTAGGATTTGGATGATACAGCTTATAGTATGAT 2271  
Db 596 ThrIleSerGlyThrLeuGlyLeuValLysAsnSerGlnAspCysCysAspAsnHisGly 615  
Qy 2272 TGGTAGTGTCTAAT---CAAAAATCAATGTCTCGAAATTTACAGTTAGGAGCTTAAGCCC 2328  
Db 616 MetPheAsnLysAspLeuGlnValProIleLeuGluLeuLysAlaThrSerAsnThr 635  
Qy 2329 CCAGCTAATGCCCATCAGATTG---ACTTAGGGAATGAGATGCCCTAAGTATGGCTAT 2385  
Db 636 ValThrThrThrAspPheSerLeuGlyThrAsnGlyTyrGlnGlnSerProTyrGlyTyr 655  
Qy 2386 CAAGGAAGCTGGAAGCTTGGTGGATCCTAATAACGCAATAATGCTCTTATCTCTG 2445  
Db 656 GlnGlyThrTyrGluPheThrIleAspThrThr-----HisThrVal 670  
Qy 2446 AAAGCTACATGCACTAAACCTGGGTATAATCTCGGCTCAGCGAGTAGCTTCTTTGGTT 2505  
Db 671 ThrGlyAsnTyrLysLysThrGlyTyrLeuProHisProGluArgLeuAlaProLeuIle 690  
Qy 2506 CCAATATGTTATGGGATCCATTTTATAGATATACGATTCGGCATTCGCAATTCAGCA 2565  
Db 691 ProAsnSerLeuTyrAlaAsnValIleAspLeuArgAlaValSerGlnAla-----Ser 708  
Qy 2566 AGTGTGATGGCCCTCTTAT---TGTCGAGGATTTGGTTTCTGGAGTTTCCGAATTC 2622  
Db 709 AlaAlaAspGlyGluAspValProGlyLysGlnLeuSerIleThrGlyIleThrAsnPhe 728  
Qy 2623 TTCTATCATGACCGCATGCTTTAGGTGAGGATATCGGTATATAGTGGGGGTATTATCC 2682  
Db 729 PheHisAlaAsnHisThrGlyAspAlaArgSerTyrArgHisMetGlyGlyTyrLeu 748  
Qy 2683 TTAGGACCAATCTCTAC-----TTTGGATCATCGATGTTTGGTGTAGCATTTTACC 2733  
Db 749 Ile-----AsnThrTyrThrArgIleThrProAspAlaAlaLeuSerLeuGlyPheGly 766

```
QY 2734 GAAGTATTGGTAGATCTAAAGATTATGAGTGCTGGTTCCTCAATCATCATGCTTGCATA 2793
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
767 GlnLeuPheThrLysSerLysAspPyrLeuValGlyHisGlyHisSerAsnValTyrPhe 786
QY 2794 CGATCCGTTTATCTATCTACCCCAACAAGCTTTTATGTGGA---TCCTATTGTTCGGAGAT 2850
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
787 AlaThrValTyrSerAsnIleThrLysSerLeuPheGlySerSerArgPhePheSerGly 806
QY 2851 GCGTTTATCCGCTAGCTACCGGTTTGGGATTCAGCATATCAAAACCTCATATACATTT 2910
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807 GlyThrSerValThrTyrSerArgSerAsnGluLysValTyrThrSerTyrThrLys 826
QY 2911 GCAGAGGAGCGATGTTCTGGGATAATAACTGCTCGCTGGAGAGATTGGAGCGGGA 2970
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
827 LeuProLysGlyArgCysSerIrpSerAsnAsnCysTrpLeuGlyGluLeuGluGlyAsn 846
QY 2971 TTACCGATTGTGATTATCCATCTAAGCTCTATTGAATGAGTGGCTCTTTCGTGCA 3030
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
847 LeuProIleThrLeuSerSerArgIleLeuAsnLeuLysGlnIleIleProPheValLys 866
QY 3031 GCTGAGTTTCTTATGCCCATCATGAATCTTTTACAGAGGAGCGCATCAAGCTCGGCA 3090
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
867 AlaGluValAlaTyrAlaThrHisGlyGlyIleGlnGluAsnThrProGluGlyArgIle 886
QY 3091 TTCAAGAGCGGACATCTCTAAATCTATCAGTTCCTGTTGGAGTGAAGTTTCATCGATGT 3150
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
887 PheGlyHisGlyHisLeuLeuAsnValAlaValProValGlyValArgPheGlyLysAsn 906
QY 3151 TCTAGTACACATCTCTAAATATAGCTTTATGCGCGCTTATCTGTGATGCTTATCGC 3210
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
907 SerHisAsnArgProAspPheTyrThrIleValAlaTyrAlaProAspValTyrArg 926
QY 3211 ACCATCTCTGGTACTGAGAACGCTCTATATCCATCAAGAGACATGGACACAGATGCC 3270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
927 HisAsnProAspCysAspThrThrLeuProIleAsnGlyAlaThrTrpThrSerIleGly 946
QY 3271 TTTCATTAGCAGACATGGAGTTGTGGTAGAGGATCTATGTATGCTCTCTCAACAGT 3330
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
947 AsnAsnLeuThrArgSerThrLeuLeuValGlnAlaSerSerHisThrSerValAsnAsp 966
QY 3331 AATATAGAATATATGGCCATGGAAGATATGATGATCGAGATGCTTCTTCGAGCTATGCT 3390
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
967 ValLeuGluIlePheGlyHisCysGlyCysAspIleArgThrSerArgGlnTyrThr 986
QY 3391 TTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
987 LeuAspIleGlySerLysLeuArgPhe 995

RESULT 5
B86547
polymorphic outer membrane protein G family [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86547
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie, T.
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: B86547
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STO>
A:Cross-references: UNIPROT:Q92896; UNIPARC:UPI000004707B; GB:BA000008; NID:98978824; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmu13
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 5.63e-86 Length: 973
Score: 1365.50 Matches: 350
Percent Similarity: 49.9% Conservative: 171
Best Local Similarity: 33.5% Mismatches: 420
```

```
Query Match: 17.3% Indels: 103
DB: 2 Gaps: 27
US-10-701-844-1 (1-4435) x B86547 (1-973)
QY 382 ATGCAACAGCTTTTCCATTAAGTCTTTCTTTCAATGATCTTAGCTTATTCTTGTGCTCT 441
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAlaProCysPheAlaSer 20
QY 442 TTAATGGGGGGGATATGCAGCAGAAATCATGTTCTCTCAAGGAATTTACATGATGGGAG 501
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
21 Thr-----AlaPheThrValGluValIleMetProSerGluAsnPheAspGlySer 37
QY 502 ACGTTAACTGATCATTTCCCTATATAGAGATCCGAGTGGAGCTACTGTTTTT 561
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 SerGlyLysIle---PheProTyrThrThrLeuSerAspProArgGlyThrLeuCysIle 56
QY 562 TCTGAGAGAGATTAAACATTAATAAATCTTGACATCTTATTGACGCTTTCCTTTAAGT 621
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 PheSerGlyAspLeuTyrIleAlaAsnLeuAspAsnAlaIleSerArgThrSerSerSer 76
QY 622 TGTTTTGGGAACCTTATTAGGAGGATTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
77 CysPheSerAsnArgAlaGlyAlaLeuGlnIleLeuGlyLysGlyValPheSerPhe 96
QY 682 GAGAACATACGACCTCTCAAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGA--- 738
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
97 LeuAsnIleArgSerSerAlaAspGlyAlaAlaIleSerSerValIleThrGlnAsnPro 116
QY 739 -----CTGTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTTCAATGTTCAATTCA 789
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 GluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePheAspAsnCysGluSer 136
QY 790 TTACTTGGCTGATCTGCTGCAACGACTAATAGGCTAGCCAGCTCCGACGACCAACA 849
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 LeuThr-----SerAspThrSerAlaSerAsn 145
QY 850 TCTACACCGTCTAATGTTACTATTATTCTAAACAGATCTTTTGTACTCAATTAATGAG 909
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 ValIleProHisAlaSerAlaIleTyrAlaThrProMetLeuPheThrAsnAsnAsp 165
QY 910 AAGTTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAC 969
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 SerIleLeuPheGlnTyrAsnArgSerAlaGlyPheGlyAlaAlaIleArgGlyThrSer 185
QY 970 TTAACGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAAATACTCTCAAGCTGAT 1029
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 IleThrIleGluAsnThrLysLysSerLeuLeuPheAsnGlyAsnGlySerIleSerAsn 205
QY 1030 GGGGAGCTTGTCAAGTAGTCACCATGTTCTCTGCTATGCTTAACGAGGCTCTATTGGC 1089
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 GlyGlyAlaLeuThrGlySerAlaAlaIleAsnLeuIleAsnAsnSerAlaProValIle 225
QY 1090 TTTTAGCGAATTTGTCAGGAGTAGAGGGGAGGAGGATTTGCTGTTCCAGGATGGGAG 1149
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
226 PheSerThrAsnAlaThrGlyIleTyrGlyAlaIleTyrLeuThr----- 241
QY 1150 CAGGAGTGTCATCATCTACTTCAACAGAGATCCAGTAGTAGTATTGTTTCCAGAAATACT 1209
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
242 ---GlyGlySerMetLeuThrSer-----GlyAsnLeuSer 252
QY 1210 GCGGTAGATTTGATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGCGAAC 1269
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 GlyValLeuPheValAsnAsnSerArgSerGlyGlyAlaIleTyrAlaAsnGlyAsn 272
QY 1270 GTTGCTTTCTCTCAATAATGAAAAACCTTTGTTTCTCAACAATGTTGCTTCTCTGTT--- 1326
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
273 ValThrPheSerAsnAsnSerAspLeuThrPheGlnAsnAsnThrAlaSerProGlnAsn 292
QY 1327 TACATTGCTTAAAGCAACCAAGTGGACAGGCTTCTTAATACGAGTAAATAATACGGA 1386
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 SerLeuProAlaProThrProProThrProAlaValThrPro---LeuLeuGly 311
```



polymorphic outer membrane protein g family - Chlamydomphila pneumoniae (strain CWL029)  
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: F72076  
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: F72076  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-973 <ARN>  
A:Cross-references: UNIPROT:Q92896; UNIPARC:UPI000004707B; GB:AE001629; GB:AE001363; NID  
A:Experimental source: strain CWL029  
C:Genetics:  
A:Gene: pmp\_13  
C:Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G

Alignment Scores: 5,63e-86 Length: 973  
Score: 1365.50 Matches: 350  
Percent Similarity: 49.9% Conservative: 171  
Best Local Similarity: 33.5% Mismatches: 420  
Query Match: 17.3% Indels: 103  
DB: 27 Gaps: 27

US-10-701-844-1 (1-4435) x F72076 (1-973)

QY 382 ATGCAACGCTTTCATAGTCTTTCTTCAATGATCTAGCTTATCTGCTCTCT 441  
DB 1 MetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAlaProCysPheAlaSer 20  
QY 442 TTAATGGGGGGGATATGACGACGAGAAATCATGCTCTCTCAAGAAATTCACGATGGGAG 501  
DB 21 Thr-----AlaPheThrValGluValIleMetProSerGluAsnPheAspGlySer 37  
QY 502 ACGTTAACTGATCATCTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
DB 38 SerGlyLysIle---PheProTyrThrThrLeuSerAspProArgGlyThrLeuCysIle 56  
QY 562 TCTGCAGGAGTAACTATAAATAATCTTGACAAATCTTATGCAAGCTTTGCTTAAAGT 621  
DB 57 PheSerGlyAspLeuTyrIleAlaAsnLeuAspAsnAlaIleSerArgThrSerSer 76  
QY 622 TGTTTGGGAACCTATTAGGAGTTTACCTGTTTATAGGAGAGACACTCGTTGACTTTC 681  
DB 77 CysPheSerAsnArgAlaGlyAlaLeuGlnIleLeuGlyLysGlyValPheSerPhe 96  
QY 682 GAGACATACGACTCTTACAAATGGGCGAGCTTAAGTAATAGCGCTGCTGATGGA--- 738  
DB 97 LeuAsnIleArgSerSerAlaAspGlyAlaAlaIleSerSerValIleThrGlnAsnPro 116  
QY 739 -----CTGTTTACTATGAGGGTTTTTAAAGAAATTTATCTTTTCCAAATTCGAATTC 789  
DB 117 GluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePheAspAsnCysGluSer 136  
QY 790 TTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACACA 849  
DB 137 LeuThr-----SerAspThrSerAlaSerAsn 145  
QY 850 TCTACCGCTCTAATGCTATTATTCTTAAACAGACTCTTTTGTGTTACTCAATAATGAG 909  
DB 146 ValIleProHisAlaSerAlaIleTyrAlaThrThrProMetLeuPheThrAsnAsnAsp 165  
QY 910 AAGTTCTCATCTAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTTAAGAGC 969  
DB 166 SerIleLeuPheGlnTyrAsnArgSerAlaGlyPheGlyAlaAlaIleArgGlyThrSer 185  
QY 970 TTAACGGTTCAAGGAATTTAGCAAGCTTTGTGCTTCCAGAAATATCTGCTCAAGCTGAT 1029  
DB 186 IleThrIleGluAsnThrLysSerLeuLeuPheAsnGlyAsnGlySerIleSerAsn 205  
QY 1030 GGGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCTACGAGGCTCTTATGCC 1089

DB 206 GlyGlyAlaLeuThrGlySerAlaAlaIleAsnLeuIleAsnAsnSerAlaProValIle 225  
QY 1090 TTTGTAGCGAATGTTGCAGGAGTAAGAGGGGAGGATTGCTGCTGTTCCAGGTGGCAG 1149  
DB 226 PheSerThrAsnAlaThrGlyIleTyrGlyAlaIleTyrLeuThr----- 241  
QY 1150 CAGGAGGTGTCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCGAAATACT 1209  
DB 242 ---GlyGlySerMetLeuThrSer-----GlyAsnLeuSer 252  
QY 1210 GCGGTAGAGTTTGTAGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAC 1269  
DB 253 GlyValLeuPheValAsnAsnSerSerArgSerGlyAlaIleTyrAlaAsnGlyAsn 272  
QY 1270 GTTGCTTCTCTCAATAAGGAAAAACCTGTTTCTCAACATGTTGCTTCTCTCTCT 1326  
DB 273 ValThrPheSerAsnAsnSerAspLeuThrPheGlnAsnAsnThrAlaSerProGlnAsn 292  
QY 1327 TACATTGCTGCTAAGCAACCAACAGTGGACAGGCTTCTTAATACGAGTAATTAATCGGA 1386  
DB 293 SerLeuProAlaProThrProProThrProAlaValThrPro---LeuLeuGly 311  
QY 1387 GATGAGGAGCTATCTTCTGTAAGATGTTGGCGCAAGCAGGATCCAACTCTGGATCA 1446  
DB 312 TyrGlyGlyAlaIlePheCysThrProAlaThrProProThrGlyVal---Ser 330  
QY 1447 GTTCTCTTGTAGGAGGAGTGTCTTCTTAGTCAATGTAGCTGCTGGGAAGG 1506  
DB 331 LeuThrIleSerGlyGluAsnSerValThrPheLeuGluAsnIleAlaSerGluGlnGly 350  
QY 1507 CGAGCTATTATATGCCAAAAGCTCTCGTGTCTAACTGTGGCCCTGTACAAATTTTAAAG 1566  
DB 351 GlyAlaLeuTyrGlyLysIleSerIleAspSerAsnLysSerThrIlePheLeuGly 370  
QY 1567 AATATGCTAATGATGCTGGAGCGATTTATTTAGGAGATCTCGAGAGCTCAGTTTATCT 1626  
DB 371 AsnThrAlaGlyLysGlyAlaIleAlaIleProGluSerGlyGluLeuSerLeuSer 390  
QY 1627 GCTGATTATGAGATTTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCT 1686  
DB 391 AlaAsnGlnGlyAspIleLeuPheAsnLysAsnLeuSerIleThrSerGly----- 407  
QY 1687 GCGGATTTAATGGCGTAATCTGTCTCCTCACAGCCATTTCGATGGGATCGGAGGAAA 1746  
DB 408 -----ThrProThrArgAsnSerIleHisPheGlyLysAspAlaLys 421  
QY 1747 ATAACGACATTAAGAGCTAAAGCGGCATCAGATCTCTTTAATGATCCATCGAGATG 1806  
DB 422 PheAlaThrLeuGlyAlaThrGlnGlyTyrThrLeuTyrPheTyrAspProIle----- 439  
QY 1807 GCMAACGGAATAACCGCCAGCGAGTCTTCCAACTTCTAAATAATTAAC----- 1857  
DB 440 ThrSerAspAspLeuSerAlaAlaSerAlaAlaThrValValValAsnProLysAla 459  
QY 1858 -----GATGGTGAAGGATACACAGGGGATATTGTTTTT----- 1890  
DB 460 SerAlaAspGly---AlaTyrSerGlyThrIleValPheSerGlyGluThrLeuThrAla 478  
QY 1891 -----GCTAATGGAACGAGTCTTTGTACCAAAATGTTACGATAGAG 1932  
DB 479 ThrGluAlaAlaThrProAlaAsnAlaThrSerThrLeuAsnGlnLysLeuGluLeuGlu 498  
QY 1933 CAAGGAAGATGTTCTTCTGTAAGGCAAAATTTATCAGTGAATTTCTCTAAGTCAGACA 1992  
DB 499 GlyGlyThrLeuAlaLeuArgAsnGlyAlaThrLeuAsnValHisAsnPheThrGlnAsp 518  
QY 1993 GGTGGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACCA 2049  
DB 519 GluLysSerValIleMetAspAlaGlyThrThr-----LeuAlaThrThrAsnGly 536  
QY 2050 CCACACAGCTCTCTGCGCTAATCAATGATCAGCTTTCCTTCCATTTGCTGCTCTT 2109

```
Db 537 AlaAsnAsnThrAspGlyAla-----lleThrLeuAsnLysLeuValilleAsnLeu 553
Qy 2110 TCTTCTTTGTAGCAAAACAATGCAAGTACGAATCTCTTACCAATCTCCAGCGCAAGAT 2169
Db 554 AspSerLeuAspGlyThrLysAla-----
Qy 2170 TCTCATCTCGAGCTATTGGT---AGCACAACTGCTGTTCTGTGTACAAATTAGTGGGCT 2226
Db 562 -----AlaValValAsnValGlnSerThrAsnGlyAlaLeuLeuThrilleSerGlyThr 578
Qy 2227 ATCTTTTTCAGGATTTCGATGATACAGCTTATGATAGTATGATGATTGGCTAGGTTCTAAT 2286
Db 579 LeuGlyLeuValLysAsnSerGlnAspCysCysAspAsnHisGlyMetPheAsnLysAsp 598
Qy 2287 ---CAAAAAATCAATGCTCGAAATTACAGTTAGGCAATAAGCCGCCAGCTAAATGCCCA 2343
Db 599 LeuGlnValProilleLeuGluLeuLysAlaThrSerAsnThrValThrThrAsp 618
Qy 2344 TCAGATTGG---ACTCTAGGGAATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAG 2400
Db 619 PheSerLeuGlyThrAsnGlyTyrglnGlnSerProfyrglyTyrglnglyThrTrpGlu 638
Qy 2401 CTTCGGTGGGATCTTAATACAGCAAAATAATGCTCCTTATCTCTGAAAGCTACATGGACT 2460
Db 639 PheThrilleAspThrThr-----HisThrValThrGlyAsnTrpLys 653
Qy 2461 AAAACTGGGTAAATCTGGGCCCTGAGCGAGTAGCTCTTTGGTTCCAAATAGTTTATGG 2520
Db 654 LysThrGlyTyrrLeuProHisProGluArgLeuAlaProLeuileProAsnSerLeuTrp 673
Qy 2521 CGATCCATTTAGATATAGTACTCGGCATTCGCATTCAGCAATTCAGCAAGTGTGGATGGCGC 2580
Db 674 AlaAsnValilleAspLysArgAlaValSerGlnAla-----SerAlaAlaAspGlyGlu 691
Qy 2581 TCTTAT---TGCAGGAGATTATGGGTTCTGGAGTTTCGAATTTCTTCTATCATGACCGC 2637
Db 692 AspValProGlyLysGlnLeuSerilleThrGlylleThrAsnPhenPheHisAlaAsnHis 711
Qy 2638 GATGCTTTAGGTCAGGATATCGGPATATTAGTGGGGTTATTCCTTAGGAGCAAACTCC 2697
Db 712 ThrGlyAspAlaArgSerTyrrArgHisMetGlyGlyGlyTyrrLeuile-----AsnThr 729
Qy 2698 TAC-----TTTGGATCATCGATGTTTGGTCTAGCATTTACCGAAGTATTGGTAGA 2748
Db 730 TyrrThrArgilleThrProAspAlaAlaLeuSerLeuGlyPheGlyGlnLeuPheThrLys 749
Qy 2749 TCTAAGATTATGATGTGTCGTTCCATCATCATGCTTGCATAGGATCGTTTATCTA 2808
Db 750 SerLysAspTyrrLeuValGlyHisGlyHisSerAsnValTyrrPheAlaThrValTyrrSer 769
Qy 2809 TCTACCCCAACAGCTTTATGTGA---TCCTATTGTTTCGGAGATGCGTTTATCCGTGCT 2865
Db 770 AsnilleThrLysSerLeuPheGlySerSerArgPhePheSerGlyGlyThrSerArgVal 789
Qy 2866 AGCTACGGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCGAGAGAGCGCAT 2925
Db 790 ThrTyrrSerArgSerAsnGlyLysValLysThrSerTyrrThrLysLeuLeuProLysGlyArg 809
Qy 2926 GTTCGTTGGGATAAATCTGCTGCGAGAGATTGCGAGCGGATTAACCGATTGCGATT 2985
Db 810 CysSerTrpSerAsnAsnCysTrpLeuGlyLysLeuGlyLysLeuProlleThrLeu 829
Qy 2986 ACTCCATCTAAGCTATTATTGAATGAGTTGCGTCTTTTCGTGCAAGCTGAGTTTCTTAT 3045
Db 830 SerSerArgilleLeuAsnLeuLysGlnilleProPheValLysAlaGluValAlaTyrr 849
Qy 3046 GCCGATCATGAATCTTTTACAGAGAGCGCATCAAGCTCGGCATTTCAGAGCGGAGAT 3105
Db 850 AlaThrHisGlyGlylleGlnGluAsnThrProGluGlyArgillePheGlyHisGlyHis 869
Qy 3106 CTCCTAAATCTATAGTCTCTCGAGTGAAGTTTGTATCGATGTTCTAGTACACATCCT 3165
Db 870 LeuLeuAsnValAlaValProValGlyValArgPheGlyLysAsnSerHisAsnArgPro 889
```

```
Qy 3166 AATAAATATAGCTTTATGGCGGCTTATATCTGTGATGCTTATCCACCACTCTCTGGTACT 3225
Db 890 AspPheTyrrThrilleValleValAlaTyrrAlaProAspValTyrrArgHisAsnProAspCys 909
Qy 3226 GAGACAACGCTCCTATCCCATCAAGAGACATGAGCAACAGATGCGCTTTTCATTTAGCAAGA 3285
Db 910 AspThrThrLeuProilleAsnGlyAlaThrTrpThrSerilleGlyAsnAsnLeuThrArg 929
Qy 3286 CATGAGCTTGTGGTTAGAGGATCTATGTATGCTTCTCTAAACAAGTAATAGAAAGTATAT 3345
Db 930 SerThrLeuLeuValGlnAlaSerSerHisThrSerValAsnAspValLeuGluilePhe 949
Qy 3346 GGCATCGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTGTAGTCGAGGAAGT 3405
Db 950 GlyHisCysGlyCysAspilleArgArgThrSerArgGlnTyrrThrLeuAspilleGlySer 969
Qy 3406 AGACTCCGGTTC 3417
Db 970 LysLeuArgPhe 973

RESULT 7
A:181723
Polymorphic membrane protein H family TC0264 [imported] - Chlamydia muridarum (strain N1)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Dec-2002
C:Accession: A81723
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684955
A:Accession: A81723
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-983 <TST>
A:Cross-references: UNIPARC:UPI0000057868; GB:AE002293; GB:AE002160; NID:g7190298; PIDN:
A:Experimental source: strain N199 (MoPn)
C:Genetics:
A:Gene: TC0264
C:Superfamily: Polymorphic membrane protein H family
```

```
Alignment Scores:
Pred. NO.: 3.95e-73 Length: 983
Score: 1179.00 Matches: 232
Percent Similarity: 78.0% Conservative: 27
Best Local Similarity: 69.9% Mismatches: 56
Query Match: 15.0% Indels: 17
DB: 2 Gaps: 3
```

US-10-701-844-1 (1-4435) x A81723 (1-983)

```
Qy 3442 GTGTAGCGATGCTTTTCTTTTGTAGATCTACATCATTTTGTGTTTGTAGCTTGTGTGT 3501
Db 1 MetLeuValMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 20
Qy 3502 TCCTATTGTTATGATTGCGAGCTCTCTCAAGTGTAAAGCCTTAATGTAAACACCTCT 3561
Db 21 SerTyrrSerTyrrGlyLeuAlaSerSerProGlnValLeuThrProAsnValillePro 40
Qy 3562 TTTAAGGAGACGATGTTTACTTGAATGGAGACTGCGCTTTGTCAATGCTATGTCAGGA 3621
Db 41 PheLysGlyAspAspilleTyrrLeuAsnGlyAspCysValPheAlaSerilleTyrrAlaGly 60
Qy 3622 GCTCAAGAAGGTTTCGATTATCTCAGCTAAATGGCGACAAATTTAAGCATTTACCGGACAAAC 3681
Db 61 AlaGluGlnGlySerilleSerAlaAsnGlyGlnAsnLeuThrilleValGlyGlnAsn 80
Qy 3682 CATACATTATCATTTACAGATTCTCAAGGGCCAGTCTTCAAAATATTATGCTTCAATTCA 3741
Db 81 HisThrLeuSerPheThrAspSerGlnGlyProAlaLeuGlnAsnCysAlaPheileSer 100
Qy 3742 GCAGGAGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCT 3801
```

DB:	2	Gaps:	34
US-10-701-844-1 (1-4435) x G86546 (1-928)			
QY	382	ATGCAACGCTCTTTTCATCAAGTCTCTTTCTTCAATGATCTAGCT---	TATCTTGTGTCG 438
Db	1	MetLysSerGlnPheSerTrpLeuValLeuSerSerThrLeuAlaCysPheThrSerCys	20
QY	439	TCCTTTAAATGGGGGGGATATGACGACGAATCATGCTTCTCAAGAAATTTACATGGG	498
Db	21	SerThrValPheAlaAla---ThrAlaGluAsnIleGlyProSerAspSerPheAspGly	39
QY	499	GAGACGTTAACTGTATCATTTCCCTATCTGTTATTAGCAGATCCGAGTGGGACTACTGTT	558
Db	40	SerThrAsnThrGlyThr-----TyThrProLysAsnThrThrThrGlyIleAspTyr	57
QY	559	TTTTCTGCAGGAGATTAAACATTAATAAAATCTTGACAAATCTATGTCAGCTTGCCTTTA	618
Db	58	ThrLeuThrGlyAspIleThrLeuGlnAsnLeuGlyAspSer---AlaAlaLeuThrLys	76
QY	619	AGTTGTTTTGGGAACTTATTAGGAGGTTTACTGTTTTTAGGAGAGACACTCGTCTGACT	678
Db	77	GlyCysPheSerAspThrThrGluSerLeuSerPheAlaGlyLysGlyTyrSerLeuSer	96
QY	679	TTGCAGAACATACGGACTCTTACAAATGGGCGACTCTTAAGTAAATAGCGCTGCTGATGGA	738
Db	97	PheLeuAsnIleLysSerSerAlaGluGlyAlaAlaLeuSer---ValThrThrAspLys	115
QY	739	CTGTGTTACTATTGAGGGTTTTAAAGAATTATCCTTTTCCAATTGCAATTCAATTACTGCC	798
Db	116	AsnLeuSerLeuThrGlyPheSerSerLeuThrPhe-----LeuAlaAla	130
QY	799	GTA CTGCTGCTGCACGACTAATAAGGTAGCCAGACTCCGACGACAACTCTACACCG	858
Db	131	ProSerSerValIleThrThrProSerGly-----	140
QY	859	TCTAATGCTACTATTTTATTCTAAACAGATCTTTTGTGTACTCAATAATGAGAAGTTCTCA	918
Db	141	---LysGlyAlaValLysCysGlyAspLeuThrPheAspAsnAsnGlyThrIleLeu	159
QY	919	TTCTATAGTAATTTAGTCTCTGGAGATGGGGGAGCTATAGATGCTAAGAGCTTACGGTT	978
Db	160	PheLysGlnAspTyrCysGluGluAsnGlyGlyAlaIleSerThrLysAsnLeuSerLeu	179
QY	979	CAAGGAATTACGACGCTTGTGTTCTTCCAAGAAATATCTGCTCAAGCTGATGGGGAGCT	103

1000

1039	QY	TGTC	AAGTAGT	CAC	CAGTTTCT	TGCT	GTAT	GGCT	TAAC	GAGGCT	TCC	TAT	TGCCTT	TTG	AGCG	1040	
195	Db	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	196	
1099	QY	AATG	TTCCAG	GAGTA	GAGGGG	AGG	GATT	CGT	CTG	CTT	CAGG	ATGG	CGG	CAG	CAGG	ATG	1100
196	Db	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	205	
1159	QY	TCAT	CATCT	ACTT	CAAC	ACAG	AGAT	CCAC	GAT	AGT	TAAGT	TTT	TCC	AGA	AACT	ACG	1160
206	Db	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	218	
1216	QY	GAGT	TTTGAT	GCGA	ACGT	AGC	CCGA	-----	-----	-----	-----	-----	-----	-----	-----	1217	
219	Db	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	238	
1273	QY	GCTT	TCCTG	AAAT	AGG	AAAA	ACCT	GTGTT	TTCT	CAAC	AAAT	GTGCT	CTCT	CCTG	TTT	ACATT	1274

	1393	GCTGTAAGCAACCAAGTGGACAGCGTTCTAATACGAGTAATAATTACGAGATGGA	1397
QY			
	252	-----SerValThrAlaThraAlaGlyAsnGly	260
Db			
	1393	GGAGCTATCTTCGTAAAGATGGTGCGCAAGCAGGATCCAATAACTCTGGATCATGTTCC	1451
QY			





## RESULT 9

G81591  
polymorphic membrane protein G family CP0303 [imported] - Chlamydomophila pneumoniae (strain G81591)  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: G81591  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: AB1500; MUID:20150255; PMID:10684935  
A:Accession: G81591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-928 <REA>  
A:Cross-references: UNIPROT:Q9RB65; UNIPARC:UPI000002FFP0; GB:AE002161; NID  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0303  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

## Alignment Scores:

Pred. No.:	9,17e-68	Length:	928
Score:	1101.00	Matches:	324
Percent Similarity:	44.5%	Conservative:	150
Best Local Similarity:	30.5%	Mismatches:	402
Query Match:	14.0%	Indels:	188
DB:	2	Gaps:	34

US-10-701-844-1 (1-4435) x G81591 (1-928)

QY	382	ATGCAACGCTTTCCATAGTCTCTTTCAATGATCTAGCT---TATCTTCTGCTGC	438
DB	1	MetLysSerGlnPheSerTrpLeuValLeuSerSerThrLeuAlaCysPheThrSerCys	20
QY	439	TCCTTAATGGGGGGATATGCGAGCAATCATGTTCTCTCAAGAAATTTACGATGGG	498
DB	21	SerThrValPheAlaAla---ThrAlaGluAsnIleGlyProSerAspSerPheAspGly	39
QY	499	GAGACGTTAACTGCTATCTCCCTATCTACTGTTATAGGAGATCCGAGTGGGACTACTGTT	558
DB	40	SerThrAsnThrGlyThr-----TyrThrProLysAsnThrThrThrGlyIleAspTyr	57
QY	559	TTTTCTGCGAGAGTTTAACTAAATAATCTTGACAAATCTATGCGCTTTGCTTTTA	618
DB	58	ThrLeuThrGlyAspIleThrLeuGlnAsnLeuGlyAspSer---AlaAlaLeuThrLys	76
QY	619	AGTTGTTTGGGAACTTATTAGGAGTTTACTGTTTATAGGAGAGACACTGTTGACT	678
DB	77	GlyCysPheSerAspThrThrGluSerLeuSerPheAlaGlyLysGlyTyrSerLeuSer	96
QY	679	TTGAGAAACATACGACTCTTCAAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGA	738
DB	97	PheLeuAsnIleLysSerSerAlaGluGlyAlaAlaLeuSer---ValThrAspLys	115
QY	739	CTGTTTACTATTAGGGTTTAAAGAAATATCTTTTCCAAATTCATTTACTTTGCC	798
DB	116	AsnLeuSerLeuThrGlyPheSerSerLeuThrPhe-----LeuAlaAla	130
QY	799	GTACTGCTGTCGACGACTAATAGGTTAGCGACTCCGACGACCAACATCTACACCG	858
DB	131	ProSerSerValIleThrThrProSerGly-----	140
QY	859	TCTAATGCTACTATTATTCTTAAACAGACTCTTTTGTCTCAATAATGAGAGTTCTCA	918
DB	141	---LysGlyAlaValLysCysGlyGlyAspLeuThrPheAspAsnAsnGlyThrIleLeu	159
QY	919	TTCTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGAGCTTAAACGGTT	978
DB	160	PheLysGlnAspTyrCysGluGluAsnGlyGlyAlaIleSerThrLysAsnLeuSerLeu	179
QY	979	CAAGGAATTAGCAAGCTTTTGCTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCT	1038

DB	180	LysAsnSerThrGlySerIleSerPheGluGlyAsnLysSerSerAla-----	195
QY	1039	TGTCAGTAGTACACGAGTTTCTCTGCTATGCTTAACGAGGCTCCTATTGCTTGTAGCG	1098
DB	195	-----	195
QY	1099	AATGTTGAGGAGTAAGAGGGGGAGGATGCTGCTTTCAGGATGGGAGGAGGATG	1158
DB	196	-----ThrGlyLysGlyGlyAlaIleCysAla-----	205
QY	1159	TCATCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACTCGC---GTA	1215
DB	206	-----ThrGlyThr-----ValAspIleThrAsnAsnThrAlaProThr	218
QY	1216	GAGTTTGTATGGAACGTAGCCGA---GTAGGAGGAGGATTTACTCTTACGGGAACGTT	1272
DB	219	LeuPheSerAsnAsnIleAlaGluAlaAlaGlyGlyAlaIleAsnSerThrGlyAsnCys	238
QY	1273	GCTTTCTCGAATAATGGAACCTTTGTTCTTCAACAATGTTGCTTCTCTCTGTTACATT	1332
DB	239	ThrIleThrGlyAsnThrSerLeuValPheSerGluAsn-----	251
QY	1333	GCTGTAAGCAACCAAGTGGACAGGCTCTTAATACGAGTAATAATTACGAGATGA	1392
DB	252	-----SerValThrAlaGlyAsnGly	260
QY	1393	GGAGCTATCTTCTGTAAGATGGTGGCAAGCAGGATCCAATACTCTGATCAGTTTCC	1452
DB	261	GlyAlaLeu-----SerGlyAspAlaAspValThr	270
QY	1453	TTTGATGAGGAGGAGTAGTTTCTTTAGTACAAATGTAGCTGCTGGGAAAGGGAGCT	1512
DB	271	IleSerGlyAsnGlnSerValThrPheSerGlyAsnGlnAlaValAlaAsnGlyGlyAla	290
QY	1513	ATTATGCAAAAGCTCTCGTTGCTAAC-----TGTGGCCCTGTACAAATTTTAAGG	1566
DB	291	IleTyrAlaLysLysLeuThrLeuAlaSerGlyGlyGlyIleSerPheSerAsn	310
QY	1567	ATATTC-----CCTAATGATGTTGGGCGATTTATTAGGAGAACTCTGGA	1611
DB	311	AsnIleValGlnGlyThrThrAlaGlyAsnGlyGlyAlaIleSerIleLeuAlaAlaGly	330
QY	1612	GAGCTCAGTTTATCTGCTGATTTAGGAGATATTATTTTCATGGGGAAT-----	1659
DB	331	GluCysSerLeuSerAlaGluAlaGlyAspIleThrPheAsnGlyAsnAlaIleValAla	350
QY	1660	---CTTAAAGAACAGCCAAAGAGATGCTGCGGATGTTTAAATGGCGTAATCTGTCCTCA	1716
DB	351	ThrThrProGlnThrThrLysArgAsnSerIleAspIle-----	363
QY	1717	CAAGCATTTCCGATGGGATCGGAGGAGGAGAAATAACGACATTAAGAGCTAAACAGGCGCAT	1776
DB	364	-----GlySerThrAlaLysIleThrAsnLeuArgAlaIleSerGlyHis	378
QY	1777	CAGATTCTCTTAAATCATCCATCCAGATCGGCAACCGAAATAACGAGCAGCGCAGTCT	1836
DB	379	SerIlePhePheTyrAspProIleThrAla-----AsnThrAlaAlaAspSer	394
QY	1837	TCCAAATCTCTTAAATAACGATGCTGAAGGA-----TACACAGGGGAT	1881
DB	395	ThrAspThrLeuAsnLeuAsnLysAlaAspAlaGlyAsnSerThrAspTyrSerGlySer	414
QY	1882	ATGTTTTTGTCT-----	1902
DB	415	IleValPheSerGlyGlyLysLeuSerGluAlaLysValAlaAspAsnLeuThr	434
QY	1903	AGTACTTGTACCAAAATGTTACGATAGCAGAGGAGGATGTTCTTCTGTGAAAAGGCA	1962
DB	435	SerThrLeuLysGlnProValThrLeuThrAlaGlyAsnLeuValLeuLysArgGlyVal	454
QY	1963	AAATTAATCAGTGAATCTCTTAAGTCAGACA---GGTGGGAGTCTGTATATGGAGCTGGG	2019

```
Db 455 ThrLeuaspThrLysGlyPheThrGlnThrAlaGlySerSerValIleMetAspAlaGly 474
Qy 2020 AGTACATGGGATTTTGTAACTCCACACCACACACACGCTCTCTGCGCTAAATCAGTTG 2079
Db 475 ThrThrLeu-----LysAlaSerThrGluGlu 483
Qy 2080 ATCAGGCTTCCAAATCGCATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2139
Db 484 ValThrLeuThrGlyLeuSerIleProValAspSerLeu----- 496
Qy 2140 AATCTCTCTACCAATCTCCAGCGCAAGATTCTCATCTCGCAGTCATTGGTAGCACAAC 2199
Db 497 -----GlyGluGlyLysLysValIleAlaLaserAlaLala 509
Qy 2200 GCTGGTCTCTGTTACAATTAGTGGGCTATCTTTTGTGAGGATTTGGATGATACAGCTTAT 2259
Db 510 SerLysAenValAlaLeuSerGlyProIleLeuLeuLeuAenGlnGlyAenAlaTyr 529
Qy 2260 GATAGGTATGATGGCTAGGTCTTAATCAAAAAATCAATGCTGCTGAATTACAG---TTA 2316
Db 530 GluAenHisAep---LeuGlyLysThrGlnAepPheSerPheValGlnLeuSerAlaLeu 548
Qy 2317 GGGACTAAGCCCCAGCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCTTAAG 2376
Db 549 GlyThrAlaThrThrAspValProAlaValProThrValAlaThrPro---ThrHis 567
Qy 2377 TATGGCTATCAAGGAAGCTGGAAGCTTCGGGGATCCTCAATACACAGCAATAATGTCCT 2436
Db 568 TyrGlyTyrGlnGlyThrTyrGlyMetThrTyrValAspAspThrAlaSerThr---Pro 586
Qy 2437 TATACTCTGAAGCTACA-----TGGACTAAACTGGGTATATATCTCGGCTGAGGGA 2490
Db 587 LysThrLysThrAlaThrLeuAlaTyrThrAenThrGlyTyrLeuProAenProGluArg 606
Qy 2491 GTAGCTCTTGGTTCCTCAATAGTTATGCGGATCCATTTAGATATACGATCTCGCAT 2550
Db 607 GlnGlyProLeuValProAenSerLeuTyrGlySerPheSerAepIleGlnAlaIleGln 626
Qy 2551 TCAGCAATTCAGCAAGTGTGTGATGGCGCTCTTATTGTCGAGGATTAATGCTTCTGGA 2610
Db 627 GlyValIleGluArgSerAlaLeuThrLeuLeuCysSerAspArgGlyPheTyrAlaAlaGly 646
Qy 2611 GTTTCGAATTTCTTCATCATGACCGCGATGCTTTAGTTCAGGGATATCGGTATATTAGT 2670
Db 647 ValAlaAenPheLeuAepLysAepLysGlyGluLysArgLysTyrArgHisLysSer 666
Qy 2671 GGGGGTTATCTCTTA---GGAGCAACTCTTACTTTGGATCATCGATGTTGGTCTAGCA 2727
Db 667 GlyGlyTyrAlaIleGlyGlyAlaAlaGlnThrCysSerGluAenLeuIleSerPheAla 686
Qy 2728 TTTACCGAAGTATTGGTAGATCTAAAGATTATGTAGTGTGCTGCTTCCAAATCATCATGCT 2787
Db 687 PheCysGlnLeuPheGlySerAspLysAspPheLeuValAlaLysAenHisThrAspThr 706
Qy 2788 TGCATAGGATCGGTTATCTATCT-----ACC 2814
Db 707 TyrAlaGlyAlaPheTyrIleGlnHisIleThrGluCysSerGlyPheIleGlyCysLeu 726
Qy 2815 CAACAGCTTTATGTTGGATCTTATTGTTGGAGATGCGTTTATC-----CGTGTAGC 2868
Db 727 LeuAepLysLeuProGlySerTyrSerHisLysProLeuValLeuGluGlnLeuAla 746
Qy 2869 TACGGGTTTGGGAATCAGCATATAAACCCTCATATACATTTTGCAGAGGAGCGCATGTT 2928
Db 747 TyrSerHisValSerAenAspLeuLysThrLysTyrThrAlaTyrProGluValLysGly 766
Qy 2929 COTTCGGATATAACTGCTGCTGCTGAGAGATGAGCGGGATTAACCGATGCTGATTACT 2988
Db 767 SerTyrGlyAenAenAlaPheAenMetMetLeuGlyAlaSerSerHisSerTyr----- 784
Qy 2989 CCATCTAGCTTAT---TTGATGATGTTGCTCTTTCGTCAGAGCTGAGTTTCTTAT 3045
Db 785 ProGluTyrLeuHisCysPheAspThrTyrAlaProTyrIleLysLeuAenLeuThrTyr 804
```

```
Qy 3046 GCCCATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGCATTTCAAGCGGACAT 3105
Db 805 IleArgGlnAspSerPheSerGluLysGlyThrGluArgSerPheAspAspSerAen 824
Qy 3106 CTCCTAAATCTATCAGTTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACATCCT 3165
Db 825 LeuPheAenLeuSerLeuProIleGlyValLysPheGluLysPheSerAspCysAenAsp 844
Qy 3166 AATAAATATAGCTTTATGGCGGCTTATATCTGTGATCTTATCGCACCATCTCTGTGACT 3225
Db 845 PheSerTyrAspLeuThrLeuSerTyrValProAspLeuIleArgAenAspProLysCys 864
Qy 3226 GAGACAACGCTCCATCCCATCAAGACATGAGCAACAGATGCTTTTATTTAGCAAGA 3285
Db 865 ThrThrAlaLeuValIleSerGlyAlaSerTyrProGluThrTyrAlaAenAenLeuAlaArg 884
Qy 3286 CATGAGCTTGTGTAGAGGATCTATGCTTATGCTTCTCTCAACAGTAATATAGAAGTATAT 3345
Db 885 GlnAlaLeuGlnValArgAlaGlySerHisTyrAlaPheSerProMetPheGluValLeu 904
Qy 3346 GGCCATCGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTGGAGTGCAGAGT 3405
Db 905 GlyGlnPheValPheGluValArgGlySerSerArgIleTyrAenValAspLeuGlyGly 924
Qy 3406 AGATCCCGGTTTC 3417
Db 925 LysPheGlnPhe 928
```

## RESULT 10

F81591  
polymorphic membrane protein G family CP0302 [imported] - Chlamydomophila pneumoniae (stra  
C; Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C; Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 15-Jun-2001  
C; Accession: F81591  
R; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
A.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A; Reference number: AB1500; MUID:20150255; PMID:10684935  
A; Accession: F81591  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-949 <REA>  
A; Cross-references: UNIPARC:UPI00001655FA; GB:AB002192; GB:AB002161; NID:g7189226; PIDN:  
A; Experimental source: strain AR39, HL cells  
C; Genetics:  
A; Gene: CP0302  
C; Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:  
Pred. No.: 3.84e-67 Length: 949  
Score: 1092.00 Matches: 316  
Percent Similarity: 42.7% Conservative: 154  
Best Local Similarity: 28.7% Mismatches: 409  
Query Match: 13.9% Indels: 222  
DB: 2 Gaps: 31

US-10-701-844-1 (1-4435) x F81591 (1-949)

```
Qy 322 ATTAATATTAAATTTTATGAAGCGGAGTAATTAATTTATCTCTCAGCTTTGTGTG 381
Db 2 IleTyrLeuPheCysPheTyrIleAspAlaAenSerSerLeuLysAenLysSerIleThr 21
Qy 382 ATGCAAGCGTCTTCCATCAAGTCTTCTTCAATGATCTAGCTTATCTTCTTTCG 435
Db 22 MetLysThrSerIleProTyrValLeuValSerValLeuAlaPheSerCysHisLeu 41
Qy 436 TGCTCTTTAAATGGGGGGGATATGACGACAGAAATCATGTTCTCAAGGAATTTACAT 495
Db 42 GlnSerLeuAlaAen-----GluGluLeuLeuSerProAspAspPheAen 57
Qy 496 GGGGAGGAGCTTAATCTGATCATTT-----CCCTATCTGTTATA 534
```





```
QY 1129 GCTGCTGTTTCAGGATGGGCGAGGAGGTGTCATCATCTACTTCAACAGAGATCCAGTA 1188
Db 202 -----
QY 1189 GTAAGTTTTTCCAGAAATACTCGG--GTAGAGTTTGTAGTGGCAAGCTAGCCCGAGTAGGA 1245
Db 203 -----IleAlaAsnAsnThrGlyTyrValArgPheLeuSerAsnIleAlaSerThrSer 220
QY 1246 GGAGGGATTACTCTACGGGAACGTTGCTTCTCTGAATTAATGGAAACCTTTGTTCTC 1305
Db 221 GlyGlyAlaIleAspAspGluGlyThrSerIleLeuSerAsnAsnLysPheLeuTyrPhe 240
QY 1306 AACAAATGTTGCTTCTCTCTGTTTACATTCGTCTAGCAACCAACAAAGTGGACAGCTTCT 1365
Db 241 -----GluGlyAsnAlaAla 245
QY 1366 AATACGAGTAATAATTACGGAGATGGAGGAGCTATCTTCTGTAGAAGAAATGTCGCAAGCA 1425
Db 246 LysThrThr-----GlyGlyAlaIle-----CysAsnThrLysAlaSerGly 259
QY 1426 GGA-----TCCAAATACTCTGGATCAGTTTCTCTTGTGAGAGGAGTA 1470
Db 260 SerProGluLeuIleIleSerAsnAsnLys-----Thr 270
QY 1471 GTTTCTTTAGTAGCAATAGCTGCTGCGAAAGGGGAGCTATTTATGCCAAAAGCTC 1530
Db 271 LeuIlePheAlaSerAsnValAlaGluThrSerGlyGlyAlaIleHisAlaLysLysLeu 290
QY 1531 TCGGTTGCTAACTGTCGCCCTGTACAAATTTTAAAGAAATATCGCTAATCAT----- 1581
Db 291 AlaLeuSerSerGlyGlyPheThrGluPheLeuArgAsnAsnValSerSerAlaThrPro 310
QY 1582 ---GGTGGAGCGCAATTATTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGANTATGA 1638
Db 311 LysGlyGlyAlaIleSerIleAspAlaSerGlyGluLeuSerLeuSerAlaGluThrGly 330
QY 1639 GATATTATTTCATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTGCCGATGTTAAT 1698
Db 331 AsnIleThrPheValArgAsnThrLeuThrThr-----GlySerThrAsp----- 346
QY 1699 GCGGTAACTGTGCTCCTCACAGCCATTCGATCGGATCGGAGCGGAAATAACGACATTA 1758
Db 347 -----ThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLysPheThrGluLeu 364
QY 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTAAATGATCCCATCGAGATGGCAACCGGAAAT 1818
Db 365 ArgAlaAlaLysAsnHisThrIlePhePheTyrAspProIle----- 378
QY 1819 AACACGCGCGAGCTTCCAAACTTCTAAATAATTAACGATGGTAGGA----- 1869
Db 379 ThrSerGluGlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeu 398
QY 1870 -----TACACGGGATATGTTTGTCT----- 1893
Db 399 AsnProTyrGlnGlyThrIleLeuPheSerGlyGluThrLeuThrAlaAspGluLeuLys 418
QY 1894 -----AATGGAACGAGTACTTGTACCAAAATGTTACGATPAGACGAGGAGGATT 1944
Db 419 ValAlaAspAsnLeuLysSerSerPheThrGlnProValSerLeuSerGlyGlyLeuLeu 438
QY 1945 GTTCTTCGTGAAAGCGCAAAATTAATCAGTGAATTCCTCTAAGTCACAGAGTGGGAGTCTG 2004
Db 439 LeuLeuGlnLysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeu 458
QY 2005 TAT---ATGAGAGCTGGAGTACATGGGATTTTGTAACTCCACAACCAACACAGCCCT 2061
Db 459 LeuGlyMetAspSerGlyThrThrLeuSerThrThrAlaGlySer----- 473
QY 2062 CCTGCCGCTAATCAGTTGATCAGCTTTCCTCAATTCGATTTGTCTCTTTCTTCTTTGTA 2121
Db 474 -----IleThrIleThrAsnLeuGlyIleAsnValAspSerLeuGly 487
```

```
QY 2122 GCAAAACAATGAGTTTACGAATCTCTTACCATAATCTCCAGCGCAAGATTCTCATCTCGCA 2181
Db 488 LeuLysGlnProVal-----Ser 493
QY 2182 GTCATTGTTAGCACAACTGCTGTTCTGTTTACAAATTAGTGGCCCTATCTTTTTCAGGAT 2241
Db 494 LeuThrAlaLysGlyAlaSerAsnLysValIleValSerGlyLysLeuAsnLeuIleAsp 513
QY 2242 TTGGATGATACAGCTTATGATAGGTATGTTGGCTAGGTTCTTAATCAAAAATAATCAATGTC 2301
Db 514 IleGluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 532
QY 2302 CTGAAATTTACAGTTTAGGGAGCTAAAGCCCGACGCTAATGCCCCATCAGATTTGACTCTAGG 2361
Db 533 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 548
QY 2362 AATGAGATGCTT-----AAGTATGGCTATCAAGGAAGCTGG 2397
Db 549 SerLeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlyGlnTrp 568
QY 2398 AAGCTTGGCTGGATCCTTAATACACGAAATAATGTCCTTATCTCTGAAAGCTACATGG 2457
Db 569 AsnValAsnIleThrThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrp 587
QY 2458 ACTAAAACCTGGGTATAATCTCGGCTCGAGCGAGTAGCTTCTTTGTTCCAAATATGTTTA 2517
Db 588 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 607
QY 2518 TGGGATCCATTTTATAGATATACGATCTGGCGATTCAGCAATTCAGCAAGTGTGGATGG 2577
Db 608 TrpGlyValPheThrAspIleArgSerLeuGlnGlnLeuValGluIleGlyAlaThrGly 627
QY 2578 CGCTCTATTGTCGAGGATTATGGGTTTCTGAGTTTCTGAAATTTCTTCTCATCATGCCGC 2637
Db 628 MetGluHisLysGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGly 647
QY 2638 GATGCTTTAGTCAGGATATCGGTATATTTAGTGGGGTATTCTTCTTAGGAGCAACTCC 2697
Db 648 AspGluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlyGlySerAla 667
QY 2698 TACTTT---GGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTATTGGTAGATCTAA 2754
Db 668 HisThrProLysAspAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 687
QY 2755 GATTATGATGTCGTCCTTCCATCATCTGTCATAGGATCCGTTTCTTCTCTATCTATACC 2814
Db 688 AspCysPheIleAlaHisAsnAsnSerArgThrTyrGlyGlyThrLeuPhePheLysHis 707
QY 2815 CAACAAGCTTTA-----TGTGGATCTCTATTGTTTCGGAGATCGC 2853
Db 708 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyArgAlaLysPheSerGluSer 727
QY 2854 TTTATC-----CGTCTAGCTACGGG 2874
Db 728 AlaIleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 747
QY 2875 TTTGGGAATCAGCATATGAACCTCATATATTCATATTCAGAGGAGAGAGATGTTCTGTTGG 2934
Db 748 HisSerAspAsnArgMetGluThrHisTyrThrSerLeuProGluSerGluGlySerTrp 767
QY 2935 GATAATAACTGTCGTCGGAGAGATTGGACGGGATTTACCGGATTTGATTTACTCCATCT 2994
Db 768 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuProPheValLeuSerAsnPro 787
QY 2995 AAGCTCTATTGAATGAGTTGCGTCTTTCGTCGAAGCTGAGTTTCTTATGCGCGATCAT 3054
Db 788 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrValSerGln 807
QY 3055 GAATCTTTTACAGAGAGGCGATCAAGCTCGGGCATTCAGAGCGGACATCTCTCTAAAT 3114
Db 808 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 827
QY 3115 CTATCAGTTCTCTTGGAGTGAAGTTTGTATCGATGTTTCTTAGTACACATCTCTAATAATAT 3174
```

```

Db      828 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerThrTyr 847
Qy      3175 AGCTTTATGGCGCTTATATCTGATGCTTATCGCACCATCTCTGGTACTGACAGCAACG 3234
Db      848 AspLeuSerGlyPhePheValSerAspValTyrArgAenAenProGlnSerThrAlaThr 867
Qy      3235 CTCCTATCCCATCAAGACACATGACACACATGCTTTCATTTAGTCAGACATGGAGTT 3294
Db      868 LeuValMetSerProAspSerTrpLysIleArgGlyGlyAenLeuSerArgGlnAlaPhe 887
Qy      3295 GTGGTTAGAGATCTATGATGCTCTCTCAACAGTATATAGAGATATAGGCATGGA 3354
Db      888 LeuLeuArgGlySerAenAenTyrValTyrAenSerAenCysGluLeuPheGlyHisTyr 907
Qy      3355 AGATATAGTATCGAGATGCTTCTCGAGGCTATGTTGAGTCGAGGAGTAGAGTCGG 3414
Db      908 AlaMetGluLeuArgGlySerSerArgAenTyrAenValAspValGlyThrLysLeuArg 927
Qy      3415 TTC 3417
Db      928 Phe 928

RESULT 12
H86546
polymorphic outer membrane protein G family [imported] - Chlamydomphila pneumoniae (strain
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86546
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: H86546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <STO>
A:Cross-references: UNIPROT:O86164; UNIPARC:UPI0000002FFEF; GB:BA000008; NID:G8978822; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp 11
C:Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 5,25e-67 Length: 928
Score: 1090.00 Matches: 311
Percent Similarity: 43.0% Conservative: 154
Best Local Similarity: 28.8% Mismatches: 394
Query Match: 13.8% Indels: 222
DB: 2 Gaps: 31

US-10-701-844-1 (1-4435) x H86546 (1-928)
Qy      382 ATGCAACGCTCTTCCATAAGTCTCTTCTTCAATGATTCAGCTTATCTTGC----- 435
Db      1 MetLysThrSerIleProTrpValLeuValSerValLeuAlaPheSerCysHisLeu 20
Qy      436 TGCTCTTTAAATGGGGGGGATATGACGACGAAATCATGTTCTTCAAGGAATTTACGAT 495
Db      21 GlnSerLeuAlaAen-----GluGluLeuLeuSerProAspAspSerPheAen 36
Qy      496 GGGGAGAGCTTAACTGTATCATTT-----CCCTATCTGTATATA 534
Db      37 GlyAenIleAspSerGlyThrPheThrProLysThrSerAlaThrThrThrSerLeuThr 56
Qy      535 GGAGATCCGAGTGGGACTACTGTTTCTGACGAGGAGTAAACATTAATAAATCTTGAC 594
Db      57 GlyAsp-----ValPhePheTyrGluProGlyLysGlyThr----- 68
Qy      595 AATCTATTGCACTTGCCTTTA-----AGTTGTTTGGGAACCTATTAGGGAGTTT 648
Db      69 -----ProLeuSerAspSerCysPheLysGlnThrThrAspAenLeu 82
```

```

Qy      649 ACTGTTTTCAGGAGGAGCAGACTCGTTGACTTTCGAGAAACATACGCGACTTCTACAAATGGG 708
Db      83 ThrPheLeuGlyAenGlyHisSerLeuThrPheGlyPheIleAspAlaGlyThrHisAla 102
Qy      709 GCAGCTCTAAGTAAATAGCGCTGCTGATGAGCTGTTTACTATTCAGGCTTTTAAAGAAATTA 768
Db      103 GlyAlaAlaAspSerThrAlaAenLysAenLeuThrPheSerGlyPheSerLeuLeu 122
Qy      769 TCCTTTTCCAAATTCGAATTCATTACTTCGCCGTACTGCTGCTGCAACGACTAATAAGGGT 828
Db      123 SerPheAspSer----- 126
Qy      829 AGCCAGACTCGACGACCAACATCTACACGCTCTAATGCTACTATTATTCTTAAACAGAT 888
Db      127 -----SerProSerThrThrValThrThrGlyGlnGlyThrLeuSerSerAlaGlyGly 144
Qy      889 CTTTGTGTACTCAATAATGAGAGCTTCTCATCTTATAGTAAATTTAGCTTCGAGAGATGGG 948
Db      145 ValAenLeuGluAenIleArgLysLeuValValAlaGlyAenPheSerThrAlaAspGly 164
Qy      949 GGAGCTATAGATGCTAAGAGCTTAAACGTTCAAGGAATTTAGCAAGCTTTGTGCTTCCAA 1008
Db      165 GlyAlaIleLysGlyAlaSerPheLeuLeuThrGlyThrSerGlyAspAlaLeuPheSer 184
Qy      1009 GAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATG 1068
Db      185 AsnAenSerSerSerThrLysGlyGlyAla----- 194
Qy      1069 GCTAACGAGGCTCTATTGCTTTGTCGCAATGTTTCAGGAGTAAAGAGGGGAGGAT 1128
Db      195 -----IleAlaThrThrAlaGlyAlaArg----- 202
Qy      1129 GCTGCTGTTTCAGGATGGGAGGAGGTGTCATCATCTTCAACAGAAAGATCCAGTA 1188
Db      202 ----- 202
Qy      1189 GTAAGTTTTCAGAAATATCTGCG---GTAGAGTTTGTATGGAGCTGAGCCGAGTAGGA 1245
Db      203 -----IleAlaAenAenThrGlyTyrValArgPheLeuSerAenIleAlaSerThrSer 220
Qy      1246 GGAGGAGTTTACTCTCTACGGGAACGTTGCTTCTGAAATTAATGGAATAACCTTGTCTC 1305
Db      221 GlyGlyAlaIleAspAspGluGlyThrSerIleLeuSerAenAenLysPheLeuTyrPhe 240
Qy      1306 AACAAATGTTGCTTCTCTGTTTACATTTGCTGCTAAGCAACCAACAGTGGAGCGCTTCT 1365
Db      241 -----GluGlyAenAlaAla 245
Qy      1366 AATACGAGTAATAATTACGGAGATGGAGAGCTATCTTCTGTAAGAATGGTGCCCAAGCA 1425
Db      246 LysThrThr-----GlyGlyAlaIle---CysAenThrLysAlaSerGly 259
Qy      1426 GGA-----TCCAATAACTCTGGATCAGTTTCTTGTGAGAGGGAGTA 1470
Db      260 SerProGluLeuIleIleSerAenAenLys-----Thr 270
Qy      1471 GTTTCTTTTAGTACAAATGATGCTGCTGGGAAAGGGGAGCTATTATTATGCCAAAAGCTC 1530
Db      271 LeuIlePheAlaSerAenValAlaGluThrSerGlyGlyAlaIleHisAlaLysLysLeu 290
Qy      1531 TCGGTGTGTAACGTGCGCCCTGTACAAATTTTAAAGGAATATCGCTTAATGAT----- 1581
Db      291 AlaLeuSerSerGlyGlyPheThrGluPheLeuArgAenAenValSerSerAlaThrPro 310
Qy      1582 ---GCTGAGCGATTTATTATTAGAGAAATCTCGAGAGCTCAGTTTATCTGCTGATTAATCGA 1638
Db      311 LysGlyGlyAlaIleSerIleAspAlaSerGlyGluLeuSerLeuSerAlaGluThrGly 330
Qy      1639 GATATTATTTTCGATGGGAATCTTAAAGAACACACCCAAAGAGATGCTGCCGATGTTAAT 1698
Db      331 AsnIleThrPheValArgAenThrLeuThrThr-----GlySerThrAsp----- 346
Qy      1699 GCGGTAACTGTGCTCTCACAGCCATTTCGATGGATCGCGGAGGGAATAACGACATTA 1758
```





## Alignment Scores:

Pred. No.: 1.59e-66 Length: 930  
Score: 1083.00 Matches: 304  
Percent Similarity: 43.3% Conservative: 160  
Best Local Similarity: 28.4% Mismatches: 406  
Query Match: 13.7% Indels: 202  
DB: 2 Gaps: 29

US-10-701-844-1 (1-4435) x D86546 (1-930)

QY 382 ATGCAACGCTCTTCCATAAGTCTCTTCTTCAATGATTCTAGCTTATTCTTGCTGCTCT 441  
DB 1 MetLysIleProLeuHisLeuLeuIleSerSerThrLeuVal---ThrProIleLeu 19  
QY 442 TTAATGGGGGGATATGCGACGAAATC---ATGGTTCCTCAAGGAATTTACGATGGG 498  
DB 20 LeuSerIleAlaThrTyrGlyAlaAspAlaSerLeuSerProThrAspSerPheAspGly 39  
QY 499 GAGCGTTAACTGATCATTTCCCTATATCTGTATAGGAGATCCGAGTGGGACTACTGTT 558  
DB 40 AlaGlyGlySerThrPheThrProLysSerThr---AlaAspAlaAsnGlyThrAsnTyr 58  
QY 559 TTTCTGCGAGGAGTTAACTATAAAATCTTGACAATTTCTATGCGAGCTTGGCTTTA 618  
DB 59 ValLeuSerGlyAsnValTyrIleAsnAspAlaGlyLysGly---ThrAlaLeuThrGly 77  
QY 619 AGTTGTTTGGGAACCTTATTAGGGAGTTTACTGTTTGGGAGAGGACACTCGTTGACT 678  
DB 78 CysCysPheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTyrSerPheSer 97  
QY 679 TTCGAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGTCTGATGGA 738  
DB 98 PheAsnThrValAspAlaGlySerAsnAlaGlyAlaAlaAspSerThrAlaAspLys 117  
QY 739 CTGTTTACTATTGAGGTTTAAAGAAATTCCTCTTTCCAAATGTCATTCATTCTGCTCC 798  
DB 118 AlaLeuThrPheThrGlyPheSerAsnLeuSerPhe 130  
QY 799 GTACTGCTGTCGCAACGACTAATAAGGGTAGCCAGACTCCGACGACAAACA----- 849  
DB 131 AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu 150  
QY 850 TCTACACCGTCTAATGGTACTATTATTTCTATAAACAGATCTTTTGTACTCAATATGAG 909  
DB 151 AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal----- 164  
QY 910 AAGTTCTCATCTATAGTAATTACTCTCTGGAGATGGGGAGCTATAGTCTAAGAGC 969  
DB 165 -----SerAsnGluAlaAsnAsnGlyLysAlaIleThrAlaLysThr 179  
QY 970 TTAACGGTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAAAGAAATPACTGCTCAAGCTGAT 1029  
DB 180 LeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLysLeu 199  
QY 1030 GGGGAGCTGTGCAAGTAGTACCAAGTTCTCTGCTATAGCTTAACGAGGCTCTATTGCC 1089  
DB 200 GlyGlyAlaIleTyrSerSerAlaAlaAspIleSerGlyAsnThrGlyGlnLeuVal 219  
QY 1090 TTTGTAGCAATGTTGCGAGGATAGAGGGGAGGATGCTGCTGTTCCAGATGGGCAG 1149  
DB 220 PheMetAsnAsn---LysGlyGluThrGlyGlyGlyAlaLeuGlyPheGlu----- 235  
QY 1150 CAGGAGGTGTCATCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAGAATACT 1209  
DB 236 -----AlaSerSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr 252  
QY 1210 GCGGTAGAGTTTGTATGGGAACGATGCCCGAGTAGGAGGAGGATTTATCTCTACGGGAAC 1269  
DB 253 AlaThrAsp----- 255  
QY 1270 GTTGCTTTCTGTAATAATGGAACCTTGTTCTTCAACAATGTTGCTTCTCTCTTTTAC 1329

DB 255 ----- 255  
QY 1330 ATTGCTCTAAGCAACCAACAGTGGAGCGCTTCTTAATACGAGTAATAATTACGAGAT 1389  
DB 256 ---AlaAla-----GlyLys 259  
QY 1390 GGAGGCTATCTCTCTAAGAAATGGTGGCAGCAGGATCCCAATNACTCTCGATCAGTT 1449  
DB 260 GlyGlyAlaIleTyrCysGluLys-----ThrGlyGluThrProThrLeu 274  
QY 1450 TCCTTCATGAGGAGGAGTAGTTTCTTTAGTAGCAATAGTCTGCTGGGAAAGGGGA 1509  
DB 275 ThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThrGlnGlyGly 294  
QY 1510 GCTATTTATGCAAAAAAGCTCTCGTGGTTGTAACCTGTGGCCTGTACAAATTTTAAAGGAT 1569  
DB 295 AlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn 314  
QY 1570 -----ATCGCTAATGATGGTGGCGGATTTATTAGGAGATCTGGAGAG 1614  
DB 315 ArgCysGlyAsnThrAlaAlaGlyLysGlyAlaIleAlaIleAlaAspSerGlySer 334  
QY 1615 CTGAGTTTATCTGCTGATTAATGAGATATTATTTTCGATGGGAATCTTAAAGAACAGCC 1674  
DB 335 LeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsn----- 349  
QY 1675 AAGAGAAATGCTCCGATGTTAATGGCGTAATCTGTCTCTCAAGCATTTCGATGGGA 1734  
DB 350 -----ThrLeuThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyrLeuGly 367  
QY 1735 TCGGGAGGAAATAACGACATTAAGAGCTAAACAGCGCATCAGATCTCTTAAATGAT 1794  
DB 368 SerSerAlaLysIleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleTyrPheTyrAsp 387  
QY 1795 CCCATCCGAGATGGCAACGGAAAT-----AACCAACCCAGCGCAG 1833  
DB 388 ProIleAlaSerAsnThrThrGlyAlaSerAspValLeuThrIleAsnGlnProAspSer 407  
QY 1834 TCTTCCAAACTTCAAATAATTAACGATGGTGAAGGATACACAGGGGATTTGTTTTGCT 1893  
DB 408 AsnSerProLeu-----AspTyrSerGlyThrIleValPheSer 420  
QY 1894 -----AATGCAAGCAGTACTTTGTAC 1914  
DB 421 GlyGlyLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys 440  
QY 1915 CAAATGTTAGCAGTAGAGCAAGAGGATTTCTTCTGCGTGAAGGCAAAATTTACGTG 1974  
DB 441 GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLysGlyAsnValGluLeuAspVal 460  
QY 1975 AATTCTCTAAGTCAGACAGGAGTCTGTATATGGAAGCTGGAGGTACATGGGATTTT 2034  
DB 461 AsnGlyPheThrGlnThrGluGlySerThrLeuLeu----- 472  
QY 2035 GTAACCTCCAAACCAACCAACAGCTCTCCCTCAATCAGTTGATCAGCTTCCCAAT 2094  
DB 473 -----MetGlnProGlyThrLysLeuLysAlaAspThrGluAlaIleSerLeuThrLys 490  
QY 2095 CTGCAATTTGCTCTTCTTCTTCTTCTTCAACCAATGCAATGCAATCTCCCTCCTACCAAT 2154  
DB 491 LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer----- 505  
QY 2155 CCTCCAGCGCAAGATTCTCATCTCGCATCTTGGTAGCAACACTGCTGGT----- 2205  
DB 506 -----IleGluThrAlaGlyAlaAsnLys 513  
QY 2206 TCTGTTACAATTAGTGGCCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGG 2265  
DB 514 ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyrGlu--- 532  
QY 2266 TATGATTGGCTAGGTTCTTAATCAAAAATCAATGTCTCTGAAATTTACAGTTAGGAGCTAAG 2325  
DB 533 -----SerHisThrIleAsn-----GlnAlaPheThrGln 542



151 AsnLeuThrAspAenGlyThrIleLeuPheSerGlnAsnVal-----164

910 AAGTTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGCGAGCTATAGATGCTAAGAGC 969

165 -----SerAenGluAlaAenAsnAsnGlyAlaIleThrAlaLysThr 179

970 TTAACGGTTCAAGAAATTTAGCAAGCTTTGTGCTTCCAGAAATACTGCTCAAGCTGAT 1029

180 LeuSerIleSerGlyAenThrSerIleThrPheThrSerAenSerAlaLysLeu 199

1030 GGGGGAGCTTGTCAAGTAGTACCAGCTTTCTCGCTATGCTACGAGGCTCCTATTGCC 1089

200 GlyIyAlaIleTySerSerAlaAlaSerIleSerGlyAenThrGlyIleLeuVal 219

1090 TTTGTACGGAATGTTGCAGAGTAAGAGGGGAGGGATTCGCTGTTACGATGGCGAG 1149

220 PheMetAenAsn---LysGlyIleThrGlyIyAlaLeuGlyPheGlu-----235

1150 CAGGAGTGTCAATCATCTTCAACAGAAAGATCCAGTAGTAAGTTTTCAGAAATACT 1209

236 -----AlaSerSerIleThrGlnAenSerLeuPhePheSerGlyAenThr 252

1210 GCGGTAGCTTTGATGGGAAGCTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC 1269

253 AlaThrAsp-----255

1270 GTTGCTTTCTGATAATAGGAAAAACCTTGTTTCTCAACAATGTTGCTTCTCCTGTTTAC 1329

255 -----255

1330 ATTGCTCTTAAGCAACCAACAGTGGGACAGGCTTCTTAATACGAGTAATAATACGGAGAT 1389

256 ---AlaAla-----GlyLys 259

1390 GGAGGACTATCTCTCTAAGATGGTGGCGGACAGGATCCAAATAACTCTGATCAGTT 1449

260 GlyIyAlaIleTyCySerGluLys-----ThrGlyIleThrProThrLeu 274

1450 TCCTTTGTATGGAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTGTGGAAAGGGGGA 1509

275 ThrIleSerGlyAenLysSerLeuThrPheAlaGluAenSerSerValThrGlnGlyGly 294

1510 GCTATTTATGCCAAAAAGCTCTCGTTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAAT 1569

295 AlaIleCyAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn 314

1570 -----ATCGCTAATGATGTGGGAGCGATTTATTATTAGAGAAATCTGGAGAG 1614

315 ArgCySerGlyAenThrAlaAlaGlyLysGlyAlaIleAlaIleAlaAspSerGlySer 334

1615 CTCAGTTTATCTGCTGATATGGAGATATATTTTCGATGGGAATCTTAAAGAACAGGCC 1674

335 LeuSerLeuSerAlaAenGlnGlyAspIleThrPheLeuGlyAen-----349

1675 AAAGAGAATGCTGCCGATGTTATTGGGGTAACTGTGCTCCTCAAGGCCATTTCGATGGGA 1734

350 -----ThrLeuThrSerThrSerAlaProThrSerThrArgAenAlaIleTyLeuGly 367

1735 TCGGGAGGGAAATTAACACATTTAAGAGCTAAAGCAGGGCATCAGATCTCTCTTTAATGAT 1794

368 SerSerAlaLysIleThrAenLeuArgAlaAlaGlnGlyGlnSerIleTyPheTyAsp 387

1795 CCCATCGAGATGGCAACCGGAAT-----AACCAAGCCAGCGCAG 1833

388 ProIleAlaSerAenThrThrGlyAlaSerAspValLeuThrIleAenGlnProAspSer 407

1834 TCTTCCAAATCTTAAAAAATTAACGATGGTGAAGGATACACAGGGGATATTGTTTTGCT 1893

408 AenSerProLeu-----AspTySerGlyThrIleValPheSer 420

1894 -----AATGGAAGCAGTACTTCTGTTAC 1914

421 GlyGluLysLeuSerAlaAspGluAlaLysAlaAlaAspAenPheThrSerIleLeuLys 440

QY	1915	CAAAATGTTACGATAGACGAAGGATATGTTCTTCGTGAAAGCAAAATATACAGTG	1977
Db	441	GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLeuLysGlyAsnValGluLeuAspVal	460
QY	1975	AATTCTTAAGTCAGACAGGTGGAGTCTGTATATGAAAGCTGGGAGTACATGGGATTTT	2034
Db	461	AsnGlyPheThrGlnThrGluGlySerThrLeuLeu	472
QY	2035	GTAATCCACCAACACCAACAGCCCTCTGCGCTAAATCAGTTGATCAGCTTCCAAT	2094
Db	473	MetGlnProGlyThrLysLeuLysAlaAspThrGluAlaIleSerLeuThrLys	490
QY	2095	CTGCATTGCTCTTCTTCTTTGAGCAAAACAATGCAGTTACGAACTCTCTACCAAT	2154
Db	491	LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer	505
QY	2155	CCTCCAGCGCAAGATTCTCATCTGCAGTCATTTGGTAGCACAACTCTCGT	2205
Db	506	IleGluThrAlaGlyAlaAsnLys	513
QY	2206	TCTGTTACAATAGTGGCGCTATCTTTTGTGAGATTTGGATGATCAGCTTATGATAGG	2265
Db	514	ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyrglu	532
QY	2366	TATGATGGCTAGGTTCTTAATCAAAAACAAATGTCCTCGAAATACAGTTAGGAGCTAAG	2325
Db	533	SerHisThrIleAsnGlnAlaPheThrGln	542
QY	2326	CCC	2361
Db	543	ProLeuValValPheThrAlaAlaThrAlaAlaSerAspIleTyrlleAspAlaLeuLeu	562
QY	2362	AATGAGATCCCTAAGTATGCGCTATCAACGAGCTGGAGCTGCG	2406
Db	563	ThrSerProValGlnThrProGluProHisTyrglyThrGlnGlyHisThrGluAlaThr	582
QY	2407	TGG	2463
Db	583	ThrAlaAspThrSerThrAlaLysSerGlyThrMetThrTrpValThr	598
QY	2464	ACTGGTATTAATCTCGGCTGACGAGTACGCTTCTTGGTTCCAAATAGTTATGCGGA	2523
Db	599	ThrGlyTyraAsnProAsnProGluuArgAlaSerValProAspSerLeuThrAla	618
QY	2524	TCCATTTTATGATATACGATTCGCGCATTCAGCAATTCACGAAATTCAGCAAGTGTGATGGGCGCTCT	2583
Db	619	SerPheThrAspIleLeuThrLeuGlnGlnIleMetThrSerGlnAlaAsnSerIleTy	638
QY	2584	TATGTCGAGATTAATGGGTTCTGGAGTTTTCGAATTTCTTCTATCATGCCGAGTCT	2643
Db	639	GlnGlnArgGlyLeuThrAlaSerGlyThrAlaAsnPhePheHisLysAspLysSerGly	658
QY	2644	TTAGTTCAGGATATCGGTATATAGTGGGGTATTCTCTAGGAGCAACTCTCTAC	2700
Db	659	ThrAsnGlnAlaPheArgHisLysSerTyrglyTyrlleValGlySerAlaGluAsp	678
QY	2701	TTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAAGATTTTGGTAGATCTAAGATTAT	2760
Db	679	PheSerGluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysAspLeu	698
QY	2761	GTAGTGTGCTTCCATCATCATGCTTGCATAGGATCCGTTTATCTATCTACCCACAA	2820
Db	699	PheIleValGluAsnThrSerHisAsnTyrlleAlaSerLeuTyrlleGlnHisArgAla	718
QY	2821	GCTTATGTCGATCTCTATTG	2856
Db	719	PheLeuGlyGlyLeuProMetProSerPheGlySerIleThrAspMetLeuLysAspIle	738
QY	2857	ATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAAACTCA	2901
Db	739	ProLeuIleLeuAsnAlaGlnLeuSerTySerTyThrLysAsnAspMetAspThrArg	758





A:Accession: C72078  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-936 >ARN  
A:Cross-references: UNIPROT:Q92898; UNIPARC:UPI000004708C; GB:AE001627; GB:AE001363; NID  
A:Experimental source: strain CML029  
C:Genetics:  
A:Gene: pmp 7  
C:Superfamily: Chlamydophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:  
Pred. No.: 1,256-63 Length: 936  
Score: 1041.00 Matches: 310  
Percent Similarity: 44.1% Conservative: 162  
Best Local Similarity: 29.0% Mismatches: 406  
Query Match: 13.2% Indels: 192  
DB: 2 Gaps: 31

US-10-701-844-1 (1-4435) x C72078 (1-936)

```
QY 382 ATGCAACGCTCTTCCATAGTCTTTCTTCAATGATCTAGCTTATTTCTGCTCT 441
|||||:|||||
Db 1 MetLysSerValSerTrpLeuPhePheSerSerIleProLeuPheSerSerLeuSer 20
QY 442 TTAATGGGGGGGATATGACAGAGAAATCATCGTT-----CCTCAAGGAATTTACGAT 495
|||:|||||:|||||:|||||:|||||
Db 21 Ile-----ValAlaAlaGluValThrLeuAspSerSerAsnAsnSerTrpAsp 36
QY 496 GGGGAGACGTTAACTGTATCTATCTTCCCTATCTTATAGGAGATCCGAGTGGGACTACT 555
|||||:|||||:|||||:|||||:|||||
Db 37 GlySerAsnGlyThrThrPheThrValPheSerThrThrAspAlaAlaGlyThrThr 56
QY 556 GTTTTTCTGCAGAGAGATTAACTTAAATAATCTTGACAAATCTATTGACAGTTTGGCT 615
|||||:|||||:|||||:|||||:|||||
Db 57 TyrSerLeuLeuSerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAla 76
QY 616 TTAAGTGTGTTTGGGAATTTATAGGAGTTTATGTTTCTAGGAGAGACACTCGTTG 675
|||||:|||||:|||||:|||||:|||||
Db 77 SerGlyCysPheLeuGluAlaGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeu 96
QY 676 ACTTTCGAGAACATACGACT--TCTACAAATGGGCGACTCTTAAGTAATACGCTGCT 732
|||||:|||||:|||||:|||||:|||||
Db 97 LysPheAlaPheIleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAla 116
QY 733 GATGGACTGTTACTATTAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTA 792
|||||:|||||:|||||:|||||:|||||
Db 117 AspLysAsnLeuLeuPheAsnAspPheSerArgLeuSerIleIleSerCysProSerLeu 136
QY 793 CTTGCCGTACTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACCAACATCT 852
|||||:|||||:|||||:|||||:|||||
Db 137 LeuLeu-----SerProThr-----141
QY 853 ACACCGCTAATGTTACTATTATTCTAAACACAGATCTTTTGTCTACTCAATAATGAGAAG 912
|||||:|||||:|||||:|||||:|||||
Db 142 -----GlyGlnCysAlaLeuLysSerValGlyAsnLeuSerLeuThrGlyAsnSerGln 159
QY 913 TTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAGAGCTTA 972
|||||:|||||:|||||:|||||:|||||
Db 160 IleIlePheThrGlnAsnPheSerSerAspAsnGlyGlyValIleAsnThrLysAsnPhe 179
QY 973 ACGTTTCAAGGAATTACAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGG 1032
|||||:|||||:|||||:|||||:|||||
Db 180 LeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAla-----195
QY 1033 GGAGCTTGTCAAGTAGTCACCAAGTCTTCTGCTATGGTACGAGGCTCTATTGCTTT 1092
|||||:|||||:|||||:|||||:|||||
Db 195 -----195
QY 1093 GTACGGAATGTTCAGGAGTAAGAGGGGAGGATGCTGCTGTCTAGGATGGGACGAG 1152
|||||:|||||:|||||:|||||:|||||
Db 196 -----PheThrGlyLysGlnGlyValValTyrAla-----206
QY 1153 GGAGTGTCTATCTACTTCAACAGAGATCCA---GTAGTAAGTATTTTCCAGAAATACT 1209
|||||:|||||:|||||:|||||:|||||
```

```
Db 207 ---ThrGlyThrIleThrIleGluAsnSerProGlyIleValSerPheSerGlnAsnLeu 225
|||:|||||:|||||:|||||:|||||
QY 1210 GCGTAGAGTTTGATGGGAACGTAGCCGAGTAGGAGGAGGATTACTCTCTACGGGAAC 1269
|||||:|||||:|||||:|||||:|||||
Db 226 Ala-----LysGlySer-----GlyGlyAlaLeuTySerThrAspAsn 238
|||||:|||||:|||||:|||||:|||||
QY 1270 GTTGCTTCTTCAATAATGGAACACTTTGTTCTTCAACAATGTTGCTTCTCTCTTTAC 1329
|||||:|||||:|||||:|||||:|||||
Db 239 CysSerIleThrAspAsnPheGlnValIlePheAspGlyAsnSerAla-----Trp 255
|||||:|||||:|||||:|||||:|||||
QY 1330 ATTGCTGTACGACCAACCAAGTGGACAGGCTTCTAATACAGTAATAATTACGGAGAT 1389
|||||:|||||:|||||:|||||:|||||
Db 256 GluAlaAlaGln-----AlaGln 261
QY 1390 GGAGCAGCTATCTCTGTAAGAATGTTGCGCAAGCAGGATCCATAACTCTCGATCAGTT 1449
|||||:|||||:|||||:|||||:|||||
Db 262 GlyGlyAlaIleCysCys-----ThrThrAspLysThrVal 274
|||||:|||||:|||||:|||||:|||||
QY 1450 TCTTTTGTATGGAGGGAGTAGTTTCTTTTAGTACAAATGTAGTCTCTGGGAAGGGGA 1509
|||||:|||||:|||||:|||||:|||||
Db 275 ThrLeuThrGlyAsnLysAsnLeuSerPheThrAsnAsnThrAlaLeuThrTyGlyGly 294
|||||:|||||:|||||:|||||:|||||
QY 1510 GCTATTATGCCAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGAT 1569
|||||:|||||:|||||:|||||:|||||
Db 295 AlaIleSerGlyLeuLysValSerIleSerAlaGlyGlyProThrLeuPheGlnSerAsn 314
|||||:|||||:|||||:|||||:|||||
QY 1570 ATCGCTATATGAT-----GTTGGAGCGATTTATTAGGAGAACTCTGGA 1611
|||||:|||||:|||||:|||||:|||||
Db 315 IleSerGlySerSerAlaGlyGlnGlyGlyAlaIleAsnIleAlaSerAlaGly 334
|||||:|||||:|||||:|||||:|||||
QY 1612 GAGCTCAGTTATCTGCTGATTTATGAGATATTTTCGATGGGAATCTTAAAGAACAA 1671
|||||:|||||:|||||:|||||:|||||
Db 335 GluLeuAlaLeuSerAlaThrSerGlyAspIleThrPheAsnAsnAsn-----350
|||||:|||||:|||||:|||||:|||||
QY 1672 GCCAAGAGAGATGCTGCCGATGTTAATGGCGTAATCTGTCTCTCACAGCCATTCGATG 1731
|||||:|||||:|||||:|||||:|||||
Db 351 -----GlnValThrAsnGlySerThrSerThrArgAsnAlaIleAsnIle 365
|||||:|||||:|||||:|||||:|||||
QY 1732 GGATCGGGAGGGAATAACAGATTAACAGCTAAAGCAGGCGCATCAGATCTCTTTAAT 1791
|||||:|||||:|||||:|||||:|||||
Db 366 IleAspThrAlaLysValThrSerIleArgAlaAlaThrGlyGlnSerIleTyPheTyr 385
|||||:|||||:|||||:|||||:|||||
QY 1792 GATCCCATCGAGATGGCAAAACGGAATAACACCGACGCGCAGCTCTTCCAACTCTTAAAA 1851
|||||:|||||:|||||:|||||:|||||
Db 386 AppProIleThr-----AsnProGlyThrAlaAlaSerThrAspThrLeuAsn 401
|||||:|||||:|||||:|||||:|||||
QY 1852 ATTAACGATGGTGAAGGA-----TACACAGGGGATATTCTTTT-----1890
|||||:|||||:|||||:|||||:|||||
Db 402 LeuAsnLeuAlaAspAlaAsnSerGluIleGluTyGlyAlaIleValPheSerGly 421
|||||:|||||:|||||:|||||:|||||
QY 1891 -----GCTAATGGAGCAGTACTTTGTACCAA 1917
|||||:|||||:|||||:|||||:|||||
Db 422 GluLysLeuSerProThrGluLysAlaIleAlaAlaAsnValThrSerThrIleArgGln 441
|||||:|||||:|||||:|||||:|||||
QY 1918 AATGTTTACCATAGACAGCAAGGAGATTGTTCTTCGTGAAAAGGCAAAATATCATGGAAT 1977
|||||:|||||:|||||:|||||:|||||
Db 442 ProAlaValLeuAlaArgGlyAspLeuValLeuArgAspGlyValThrValThrPheLys 461
|||||:|||||:|||||:|||||:|||||
QY 1978 TCTCTAAGTCAGACAGCTGGGAGT---CTGTATATGGAAGCTGGGAGTACATCGGATTTT 2034
|||||:|||||:|||||:|||||:|||||
Db 462 AspLeuThrGlnSerProGlySerArgIleLeuMetAspGlyGlyThrThrLeuSer---480
|||||:|||||:|||||:|||||:|||||
QY 2035 GTAATCTCCACACCAACCAACAGCAGCTCTGCGCTAATCAGTTGTATCAGCTTCCAAAT 2094
|||||:|||||:|||||:|||||:|||||
Db 481 -----AlLysGluAlaAsnLeuSerLeuAsnGly 490
|||||:|||||:|||||:|||||:|||||
QY 2095 CTCGATTGTCTCTTTCTTCTTTTAGCAAAACAATGCAGTTAGCAATCTCTCTACCAAT 2154
|||||:|||||:|||||:|||||:|||||
Db 491 LeuAlaValAsnLeuSerSerLeuAspGlyThrAsnLysAla-----504
|||||:|||||:|||||:|||||:|||||
QY 2155 CCTCCAGCGCAAGATTCTCATCTCTCCTGAGTCTGTTAGTGGTAGCAACAATGCTGTTCTGTTACA 2214
|||||:|||||:|||||:|||||:|||||
```



D	b	505	- - - - - AlaleuLysThrGluAlaAAspLysAsnIleSer	516
Q	y	2215	ATTAGTGGCCCTACTCTTTTGAGGATTTGGATGATACAGCTTATGATGATGATGATGCG	2274
D	b	517	LseUerGlyThrIleAlaLeuIleAspThrGluGlySerPheTyrgluAenHisAen---	535
Q	y	2275	CTAGGTCTTAATCAAAAATCAATGTGCTCTGAATTCACAGTTAGGAGCATAGCCCCCAGCT	2334
D	b	536	LeuLysSerAlaSerThrtYrProLeuLeuGluLeu-----ThrThralaGlyAla	552
Q	y	2335	AAT-----GCCCCATCAGATTGTGACTCTAGGGAATGAGATCCCTAAAG	2376
D	b	553	AsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGlnGluProGluThrHis	572
Q	y	2377	TATGGCTATCAAGGAAGCTGAAGCTGGCGGGATCCTAATAACAGCAAATAATGGTCCT	2436
D	b	573	TyrGlyTyrGlnGlyAsnTrpGlnLeuSerTrpAlaAsnAlaThrSerSerLys-----	590
Q	y	2437	TATACTCTGAAGCTACATGCAGTAAACCTGGGTATATCTCGGCCCTCAGCGAGTAGT	2496
D	b	591	--IleGlySerIleAsnTrpThrArgThrGlyTyrIleProSerProGluArgLysSer	609
Q	y	2497	TCCTTGGTTCAAATAGTTATGGGATCCATTTTAGNATATACATCTCGGCATTCAGCA	2556
D	b	610	AsnLeuProLeuAsnSerLeutrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu	629
Q	y	2557	ATCCAAGCAAGTGTGGATGGCGCTCTTATTGTGAGGATTATCGTTTCTGGAGTTTCG	2616
D	b	630	IleGluThrLysSerSerGlydUpProPheGluArgGluLeutrpLeuSerGlyIleAla	649
Q	y	2617	AATTCTCTCATCATGACCGCGATGCTTTAGTCCAGGATATCGGTATATTATAGTGGGGT	2676
D	b	650	AsnPhePheTyArgAspSerWetProthrArgHisGlyPheArgHisIleSerGlyGly	669
Q	y	2677	TATTCTTAGAGACAACCTCCTACTTT---GGATCATCGATGTTTTGGTCTAGCATTTACC	2733
D	b	670	TyrAlaLeuGlyIleThrAlaThrProAlaGluAspGlnLeuThrPheAlaPheCys	689
Q	y	2734	GAAGTATTGTGTAGATCTAAAGATTATGTAGTGTGTGCTTCCATCATCATGCTTGCATA	2793
D	b	690	GlnLeuPheAlaaGspArgAsnHisIleThrGlyLysAsnHisGlyAspThrTyrgly	709
Q	y	2794	GGATCCGTTTATCTATCTACCCACAAGCTTTA-----TGTCGATCCTATTCTTCGGA	2847
D	b	710	AlaSerLeuTyrrPheHisHieThrGluGlyLeuPheaspIleAlaAenPheLeutrpGly	729
Q	y	2848	GATCGGTTT-----	2856
D	b	730	LysAlaThrArgAlaProTrpValLeuSerGluIleSerGlnIlelleProLeuSerPhe	749
Q	y	2857	---ATCCGTGCTAGCTACGGTGTGGGAATCAGCATATGAAACCTCATATACATTTGCA	2913
D	b	750	AspAlaLysPheSertyrLeuHisThrAspAsnHisMetLysThrTyrrTyrrh-----	767
Q	y	2914	CAGGAGACGATGTTTCGT-----TGGGATATAATACCTCTCGGCTCGAGAGATGGACG	2967
D	b	768	AspAsnSerIlelleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAla	787
Q	y	2968	GGATTACGATGTGATPATCTCCATCTAAGCTCTATTGGAAGTAGTTCGGTCTTCGTG	3027
D	b	788	SerLeuProPheValIleSerValprotyrLeu---LeuLysGluValGluProPheVal	806
Q	y	3028	CAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGG	3087
D	b	807	LysValGlnTyrrIleTyrrAlaHisGlnGlnAspPheTyrrGluA-ghisAlaGluGlyArg	826
Q	y	3088	GCATTCAAGACGGCACATCTCTCAATCTATCATCTGTTGTGGAGTCAAGTTTCATCGA	3147
D	b	827	AlaPheAsnLysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArg	846
Q	y	3148	TGTTCTAGTACACATCTCAATAATATAGCTTTATGCGCGCTTATATCTGTGATGCTTAT	3207
D	b	847	AspSerLysSerGluLysGlyThrTyrrAspLeuThrLeuMetTyrrIleLeuAspAlaTyrr	866

QY	3208	CGCACCATCTCTGGTACTTGAGACACACGCTCTCTATCCCATCAAGACAGCATGGACACACGAT	3267
Db	867	ArgAenProLysCysGlnThrSerLeuLeuAlaSerAspAlaSerMetAlaTyr	886
QY	3268	GCCTTTCATTAGCAGACATGAGGTGTGGTTAGAGGATCTATGTATGCTCTCTTAACA	3327
Db	887	GlyThrAsnLeuAlaArgGlnGlyPheSerValArgAlaAsnHisPheGlnValAsn	906
QY	3328	AGTAATATAGAGTATATGGCCATGGAGATATAGATATCGAGATGCTTCTCGAGGCTAT	3387
Db	907	ProHisMetGluIlePheGlyGlnPheAlaPheGluValArgSerSerAsnTyr	926
QY	3388	GGTTTGATGCAGCAAGTAGATCCGGTTC	3417
Db	927	AsnThrAsnLeuGlySerLysPheCysPhe	936
RESULT 17			
B81591			
polymorphic membrane protein G family CP0308 [imported] - Chlamydomophila pneumoniae			
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae			
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004			
C:Accession: B81591			
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.;			
C: Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.;			
Nucleic Acids Res. 28, 1397-1406, 2000			
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae			
A:Reference number: AB1500; MUID:20150255; PMID:10684935			
A:Accession: B81591			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-936 <REA>			
A:Cross-references: UNIPROT:Q92898; UNIPARC:UP10000131CED; GB:AE002193; GB:AE000			
A:Experimental source: strain AR39, HL cells			
C:Genetics:			
A:Gene: CP0308			
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G			
Alignment Scores:			
Pred. No.:		1.46e-63	Length: 936
Score:		1040.00	Matches: 310
Percent Similarity:		44.1%	Conservative: 162
Best Local Similarity:		29.0%	Mismatches: 406
Query Match:		13.2%	Indels: 192
DB:		2	Gaps: 31
US-10-701-844-1 (1-4435) x B81591 (1-936)			
QY	382	ATGCAGACGCTCTTCCATAAGTCTTCTTTCAATGATCTAGCTTATCTCTGCTGCTCT	441
Db	1	MetLysSerSerValSerTrpLeuPhePheSerSerIleProLeuPheSerSerLeuSer	20
QY	442	TTAAATGGGGGGGATATGCAGCAGAAATCATGGTT-----CCTCAAGAAATTTACGAT	495
Db	21	Ile-----ValAlaAlaGluValThrLeuAspSerSerAsnSerTyrAsp	36
QY	496	GGGGAGAGCTTAACCTGCTATCATTTCCCTATCTAGTATAGAGATCCGAGTGGGACTACT	555
Db	37	GlySerAsnGlyThrThrPheThrValPheSerThrThrAspAlaAlaGlyThrThr	56
QY	556	GTCTTTTCGCGAGAGAGTAAACATTAATAAATCTTCACAAATCTATTCGACGCTTGGCTCT	615
Db	57	TyrSerLeuLeuSerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAla	76
QY	616	TTAAGTTGTTTGGAACTTATTAGGAGGTTTTACTCTTTTAGGAGAGACACTCGCTTG	675
Db	77	SerGlyCysPheLeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeu	96
QY	676	ACTTTCGAGACATACGGACT---TCTCAAAATGGGCGAGCTCTTAAGTATATAGCGTCTG	732
Db	97	LysPheAlaPheIleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAla	116
QY	733	GATGACCTGTTTACTATTATGAGGGTGTAAAGAAATTTATCTTTTCCAAATTCGAATTCATTA	792















Db 629 ValAlaThrLysValArgGlnSerGlnGluThrArgGlyIleTrpCysGluGlyIleSer 648  
 QY 2617 AATTTCTTCTATCATGACCGCGATGCTTTAGTGCAGGATATCGGTATATAGTGGGCT 2676  
 Db 649 AsnPhePheHisLysAspSerThrLysIleAsnLysGlyPheArgHisIleSerAlaGly 668  
 QY 2677 TATTCCTTAGGCAACCTCTACTTTGGATCA---TCGATGTTTGGTCTAGCATTTACC 2733  
 Db 669 TyrValValGlyAlaThrThrThrLeuAlaSerAspAsnLeuIleThrAlaAlaPheCys 688  
 QY 2734 GAAGTATTTGGTAGATCTAAAGATATAGTAGTGTGCTTCCATCATCATGCTTCCATA 2793  
 Db 689 GlnLeuPheGlyLysAspArgAspHisPheIleAsnLysAsnArgAlaSerAlaTyrAla 708  
 QY 2794 GGNATCGTTTATCTATCTACCCCAACAGCTTTATGTGGA-----TCCTATTGG 2841  
 Db 709 AlaSerLeuHisLeuGlnHisLeuAlaThrLeuSerSerProSerLeuLeuArgTyrLeu 728  
 QY 2842 TTCGGA-----GATCGGTTTATCGTGCTAGTACGGG 2874  
 Db 729 ProGlySerGluSerGluGlnProValLeuPheAspAlaGlnIle-----SerTyrIle 746  
 QY 2875 TTTGGGAATCAGCATATGAACCTCATATACATTTTCAGAGAGAGCGGATGTCGTGG 2934  
 Db 747 TyrSerLysAsnThrMetLysThrTyrThrGlnAlaProLysGlyGluSerSerTrp 766  
 QY 2935 GATAATAACGTCTGGTGGAGAGATGGAGCGGATACCGATTGGATTGATTCTCCATCT 2994  
 Db 767 TyrAsnAspGlyCysAlaLeuGluLeuAlaSerSerLeuProHisThrAlaLeuSerHis 786  
 QY 2995 AAGCTCTATTGATGATGTTGCTCTTCGTCGAAGCTGAGTTTCTTATGCGCATCAT 3054  
 Db 787 GluGlyLeuPheHisAlaTyrPheProPheIleLysValGluAlaSerTyrIleHisGln 806  
 QY 3055 GAATCTTTTACAGAGGAAGCGATCAA---GCTCGGCAATTCAGAGCGGACATCTCCTA 3111  
 Db 807 AspSerPheLysGluArgAsnThrThrLeuValArgSerPheAspSerGlyAspLeuIle 826  
 QY 3112 AATCTATCACTGTTGTTGAGTGAAGTTGATGATGCTTCTAGTACACATCTCTATAAA 3171  
 Db 827 AsnValSerValProIleGlyIleThrPheGluArgPheSerArgAsnGluArgAlaSer 846  
 QY 3172 TATAGCTTTATGCGGCTTATCTGATGCTTATCGCACCCTCTCTGCTACTGAGACA 3231  
 Db 847 TyrGluAlaThrValIleTyrValAlaAspValTyrArgLysAsnProAspCysThrThr 866  
 QY 3232 ACGCTCTATCCCATCAAGAGACATGACACACAGATGCCTTTCATTTAGCAAGACATGGA 3291  
 Db 867 AlaLeuLeuIleAsnAsnThrSerTrpLysThrThrGlyThrAsnLeuSerArgGlnAla 886  
 QY 3292 GTTGTTGTTAGAGATCTATGATGCTTCTTCAAGTATATAGATATATGCGCAT 3351  
 Db 887 GlyIleGlyArgAlaGlyIlePheTyrAlaPheSerProAsnLeuGluValThrSerAsn 906  
 QY 3352 GGAAGATATGATATCCAGATGCTTCTCGAGGCTATGCTTTGATGAGCAGAGTAGATC 3411  
 Db 907 LeuSerMetGluIleArgGlySerSerArgSerTyrAsnAlaAspLeuGlyLysPhe 926  
 QY 3412 CGGTTCT 3417  
 Db 927 GlnPhe 928

RESULT 21  
 B86546  
 polypeptide outer membrane protein G/I family [imported] - Chlamydomonas pneumoniae (str  
 C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: B86546  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ito  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349; PMID:10871362  
 A:Accession: B86546

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1276 <STO>  
 A:Cross-references: UNIPROT:Q92899; UNIPARC:UPI0000131CEC; GB:BA000008; MID:98978816; P:  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: pmp\_6  
 Alignment Scores:  
 Pred. No.: 8 25e-62 Length: 1276  
 Score: 1015.00 Matches: 342  
 Percent Similarity: 37.9% Conservatives: 152  
 Best Local Similarity: 26.3% Mismatches: 421  
 Query Match: 12.9% Indels: 388  
 DB: 2 Gaps: 38  
 US-10-701-844-1 (1-4435) x B86546 (1-1276)  
 QY 493 GATGGGAGACGTTAACTGTATCTATCTTCCCTATCTATCTTATAGGAGATCCGAGTGGGACT 552  
 Db 36 AsnGlySerSerGlySerAlaAlaPheThrAlaLysGluThrSerAspAlaSerGlyThr 55  
 QY 553 ACTGTTTCTGCGAGGAGATTAACATTTAAAATCTTGACATTTCTATTGCGAGTTTG 612  
 Db 56 ThrTyrThrLeuThrSerAspValSerIleThrAsnVal---SerAlaIleThrProAla 74  
 QY 613 CCTTTAAGTTGTTTGGGAACCTTATTAGGAGATTTTACTGTTTATAGGAGAGACACTCG 672  
 Db 75 AspLysSerCysPheThrAsnThrGlyGlyAlaLeuSerPheValGlyAlaAspHisSer 94  
 QY 673 TTGACTTTTCAGAACATCAGGACTTCTACAAATGGGCGAGCTCTAAGTAAATACGCGTGT 732  
 Db 95 LeuValLeuGlnThrIleAlaLeuThrHisAspGlyAlaAlaIleAsnAsnThrAsnThr 114  
 QY 733 GATGAGCTGTTTACTATTGAGGTTTAAAGAAATATCTTCTTCCATTTGCAATTCATT 792  
 Db 115 ---AlaLeuSerPheSerGlyPheSerSerLeu 124  
 QY 793 CTT----- 795  
 Db 125 LeuIleAspSerAlaProAlaThrGlyThrSerGlyLysGlyAlaIleCysValThr 144  
 QY 795 --- 795  
 Db 145 AsnThrGluGlyGlyThrAlaThrPheThrAspAsnAlaSerValThrLeuGlnLysAsn 164  
 QY 795 --- 795  
 Db 165 ThrSerGluLysAspGlyAlaAlaValSerAlaTyrSerIleAspLeuAlaLysThrThr 184  
 QY 796 ---GCCGTACTGCTGCT 810  
 Db 185 ThrAlaAlaLeuLeuAspGlnAsnThrSerThrLysAsnGlyGlyAlaLeuCysSerThr 204  
 QY 811 GCAACGACT-----AATAAGGTTAGCCAGCT---CCGACGACCAATCTACA 855  
 Db 205 AlaAsnThrValGlnGlnSerGlyThrValThrPheSerSerAsnThrAlaThr 224  
 QY 856 CCGTCTAATGGTACTATTATTCT----- 879  
 Db 225 AspLysGlyGlyIleTyrSerLysGluLysAspSerThrLeuAspAlaAsnThrGly 244  
 QY 879 --- 879  
 Db 245 ValValThrPheLysSerAsnThrAlaLysThrGlyAlaTrpSerSerAspAsn 264  
 QY 879 --- 879  
 Db 265 LeuAlaLeuThrGlyAsnThrGlnValLeuPheGlnGluAsnLysThrThrGlySerAla 284  
 QY 879 --- 879  
 Db 285 AlaGlnAlaAsnAsnProGluGlyCysGlyAlaIleCysCysTyrLeuAlaThrAla 304

QY 890 -----AAAACAGATCTTTTGTACTCAATATGAGAGGTTCTCTCTTCTAGTAATTTA 933  
Dbb||||| :||:||||:|  
305 ThrAspLysThrGlyLeuAlaIleSerGlnAenGlnGluMetSerPheThrSerAsnThr 324  
QY 934 GTCTCTGGAGATGGGGAGCTATAGACTCTAAGAGCTTAAACGGTTCAAGGAATAGCAAG 993  
Dbb||||| :||:||||:|  
325 ThrThrAlaAenGlyGlyAlaIleTyrAlaThrLysCysThrLeuAenGlyAsnThrThr 344  
QY 994 CTTTGTGTCTCCAAAGAAATCTCTCAAGCTGAT---GGGGAGCTTGTCAAGTAGTC 1050  
Dbb||||| :||:||||:|  
345 Leu---ThrPheAspGlnAenThrAlaThrAlaGlyCysGlyGlyAlaIleTyrThrGlu 363  
QY 1051 ACC---AGTTCTCTGCTATGGCTAACAGGCTCTATTGCTTTGTAGCGAATGTTGCA 1107  
Dbb||||| :||:||||:|  
364 ThrGluAenPheSerLeuLysGlySerThrGlyThrValThrPheSerThrAsnThrAla 383  
QY 1108 GGAGTAAGGGGGAGGGATT----- 1128  
Dbb||||| :||:||||:|  
384 LysThr---GlyGlyAlaLeuTyrSerLysGlyAenSerSerLeuThrGlyAsnThrAsn 402  
QY 1129 -----GCTGCTGTTCAGGATGGG 1146  
Dbb||||| :||:||||:|  
403 LeuLeuPheSerGlyAsnLysAlaThrGlyProSerAsnSerSerAlaAenGlnGluGly 422  
QY 1147 CAGCAGGGAGTG-----TCATCATCTACTTCAACA----- 1176  
Dbb||||| :||:||||:|  
423 CysGlyGlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysLysGlyLeu 442  
QY 1177 -----GAAGATCCAGTAGTAGTTTCCAGAAATCTGCG----- 1212  
Dbb||||| :||:||||:|  
443 TrpIleGluAenGluAenValSerLeuSerGlyAenThrAlaThrValSerGlyGly 462  
QY 1213 -----GTAGAGTTTGATGGGAAC 1230  
Dbb||||| :||:||||:|  
463 AlaIleTyrAlaThrLysCysAlaLeuHisGlyAenThrThrLeuThrPheAspGlyAen 482  
QY 1231 GTAGCCCGA---GTAGGAGGAGGATTTAC----- 1257  
Dbb||||| :||:||||:|  
483 ThrAlaGluThrAlaGlyAlaIleTyrThrGluThrGluAenPheThrLeuThrGly 502  
QY 1258 TCCTACGGAGAGTTGCTTCTCGAATAAT----- 1287  
Dbb||||| :||:||||:|  
503 SerThrGlyThrValThrPheSerThrAsnThrAlaLysThrAlaGlyAlaLeuHisThr 522  
QY 1288 ---GGAAAAACCTTGTTCTCAACAATGTTGCTTCTCTGTTTACATTTGCTGTAAGCAA 1344  
Dbb||||| :||:||||:|  
523 LysGlyAenThrSerPheThrLysAenLysAla-----LeuValPheSerGlyAen 539  
QY 1345 CCAACAGTGGACAGCTTCTTAATACGAGTAATAATTACGAGAGTGGAGGACTATCTTC 1404  
Dbb||||| :||:||||:|  
540 SerAlaThrAlaThrAlaThrThrThrAspGlnGluGlyCysGlyGlyAlaIleLeu 559  
QY 1405 TGT-----AAGAAATGGTGG 1419  
Dbb||||| :||:||||:|  
560 CysAenIleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAenGluSer 579  
QY 1420 CAAGCAGGATCCAATAACTCTCGATCAGTTTCC----- 1452  
Dbb||||| :||:||||:|  
580 LeuSerPheIleAenAenThrAlaLysArgSerGlyGlyIleTyrAlaProLysCys 599  
QY 1453 ---TTTGATGGAGAGGAGTAGTTTTCTTTTAGTACCAATGTAGCTGCTGGGAAAGGGGA 1509  
Dbb||||| :||:||||:|  
600 ValIleSerGlySerGluSerIleAenPheAspGlyAenThrAlaGluThrSerGlyGly 619  
QY 1510 GCTATTTATGCCAAAAGCTCTCGTTGCTTAAGTGGCCCTGTACAAATTTTAGGAAT 1569  
Dbb||||| :||:||||:|  
620 AlaIleTyrSerLysAenLeuSerIleThrAlaAenGlyProValSerPheThrAenAen 639  
QY 1570 ATCGCTTAAGTGGTGGAGGATTTATTAGGAGAACTGGAGAGCTCAGTTTATCTGCT 1629  
Dbb||||| :||:||||:|  
640 SerGlyGlyLysGlyAlaIleTyrIleAlaAenPheSerGlyGluLeuSerLeuGluAla 659

QY 1630 GATTATGAGATATTATTTTCGATGGGAATCTTAAAGAACAGCCAGAGAAATGCTGCC 1689  
Dbb||||| :||:||||:|  
660 IleAspGlyAspIleThrPheSerGlyAen-----ArgAlaThrGluGlyThrSer 676  
QY 1690 GATGTTAATGGCGTAACGTGCTCTCAAGCCATTTTCGATGGATCGGAGGGGAAATA 1749  
Dbb||||| :||:||||:|  
677 ThrProAen-----SerIleHisLeuGlyAlaGlyAlaLysIle 689  
QY 1750 ACGACATTAAGAGCTAAAGAGGGCATCAGATTCTCTTTTAATGATCCCATCGAGTG--- 1806  
Dbb||||| :||:||||:|  
690 ThrLysLeuAlaAlaAlaProGlyHisThrIleTyrPheTyrAspProIleThrMetGlu 709  
QY 1807 -----CCAAACGGA----- 1815  
Dbb||||| :||:||||:|  
710 AlaProAenSerGlyGlyThrIleGluGluLeuValIleAenProValValLysAlaIle 729  
QY 1816 -----AATAACCCAGCCAGCGCAGCTCTCCCAAACTTCTTAAAAATTAAC 1857  
Dbb||||| :||:||||:|  
730 ValProProGlnProLysAenGlyProIleAlaSerValProValValProValAla 749  
QY 1858 GATGGTGAAGATACACAGGGGATATTGTTTTT----- 1890  
Dbb||||| :||:||||:|  
750 ProAlaAenProAenThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAsp 769  
QY 1891 -----GCTAATGGAGCAGTACTTTGTACCAAAATGTTACGATAGACNAGGA 1938  
Dbb||||| :||:||||:|  
770 AlaSerIleProAlaAenThrThrThrIleLeuAenGlnLysIleAenLeuAlaGlyGly 789  
QY 1939 AGGATGTTCTTCTCGTGAAGGCAAAATTTATCAGTGAATTTCTTAAGTCACACAGGTGGG 1998  
Dbb||||| :||:||||:|  
790 AsnValValLeuLysGlyGlyAlaThrLeuGlnValTyrSerPheThrGlnGlnProAsp 809  
QY 1999 AGT---CTGTATATGGAAGCTGGAGTACATGGGATTTTGTAACTCCACACACCACAA 2055  
Dbb||||| :||:||||:|  
810 SerThrValPheMetAspAlaGlyThrThrLeuGluThrThrThrAsn----- 826  
QY 2056 CAGCTCTCGCGCTAATCAGTTCAATCAGCTTTCGAATCTGCATTTGTCTCTTCTCTCT 2115  
Dbb||||| :||:||||:|  
827 -----AenThrAspGlySerIleAspLeuLysAenLeuSerValAenLeuAspAla 843  
QY 2116 TTGTTAGCAACAATGCAAGTTACGAATCTCTACCAATCTCCAGCGCAAGATTTCTCAT 2175  
Dbb||||| :||:||||:|  
844 LeuAspGlyLysArgMetIleThr----- 851  
QY 2176 CTGCGAGTCATGTGTAGCACAACTGCTGTTCTGTATCAATATGATGGCCCTATCTTTT 2235  
Dbb||||| :||:||||:|  
852 -----IleAlaValAenSerThrSerGlyGlyLeuLysIleSerGlyAspLeuLysPhe 869  
QY 2236 GAGGATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAATCAAAAATC 2295  
Dbb||||| :||:||||:|  
870 HisAenAenGluGlySerPheTyrAspAen-----ProGlyLeuLysAlaAenLeu 886  
QY 2296 AATGCTCTGAAATACAGTTA---GGGACTAAGCCCCCAGCTAAT----- 2337  
Dbb||||| :||:||||:|  
887 AsnLeuProPheLeuAenLeuSerSerThrSerGlyThrValAenLeuAspAspPheAen 906  
QY 2338 ---GCCCATCAGATTTTGACTCTAGGGAATGAGATGCTTAAGTATGCTATCAAGGAAGC 2394  
Dbb||||| :||:||||:|  
907 ProIleProSerSerMetAla-----AlaProAspTyrGlyTyrGlnGlySer 922  
QY 2395 TGGAGCTGCTGGTGGATCTTAATACACCAATTAATGCTCTTACTCTGAAAGCTACA 2454  
Dbb||||| :||:||||:|  
923 TrpThrLeuVal-----ProLysValGlyAlaGlyLysValThrLeuValAlaGlu 940  
QY 2455 TGGACTAAAACCTGGGTATAATCTGGGCTCAGCGAGTAGCTCTTTGGTTCCAAATAGT 2514  
Dbb||||| :||:||||:|  
941 TrpGlnAlaLeuGlyTyrThrProLysProGluLeuArgAlaThrLeuValProAenSer 960  
QY 2515 TTATGGGATCCATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGAT 2574  
Dbb||||| :||:||||:|  
961 LeuTrpAenAlaTyrValAenIleHisSerIleGlnGlnGluIleAlaThrAlaMetSer 980  
QY 2575 GGGCGCTCTTATGTCGAGGATTAATGGTTCTCGAGTTTCGAATTTCTTCTATCATGAC 2634



Db 285 AlaGlnAlaAsnAsnProGluGlyCysGlyGlyAlaIleCysCysThyLeuAlaThrAla 304  
QY 880 -----AAAACAGATCTTTGTTACTCAATATGAGAAAGTTCTCATCTATAGTAATTTA 933  
Db 305 ThrAspLysThrGlyLeuAlaIleSerGlnAsnGlnGluMetSerPheThrSerAsnThr 324  
QY 934 GTCTCTGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAGCGTTCAAGGAAATAGCAAG 993  
Db 325 ThrThrAlaAsnGlyGlyAlaIleTyAlaThrLysCysThrLeuAspGlyAsnThrThr 344  
QY 994 CTTTGTGCTCCRAAGAAATCTCTCAAGCTGAT---GGGGAGCTTGTCAGTAGTC 1050  
Db 345 Leu---ThrPheAspGlnAsnThrAlaThrAlaGlyCysGlyGlyAlaIleTyThrGlu 363  
QY 1051 ACC---AGTTTCTCTGCTATGCTTAACAGGCTCTATTGCTTTGTAGCGAATGTCGA 1107  
Db 364 ThrGluAspPheSerLeuLysGlySerThrGlyThrValThrPheSerThrAsnThrAla 383  
QY 1108 GGAGTAAGAGGGGAGGGATT----- 1128  
Db 384 LysThr---GlyGlyAlaLeuTySerLysGlyAsnSerSerLeuThrGlyAsnThrAsn 402  
QY 1129 -----GCTGCTGTTCCAGGATGG 1146  
Db 403 LeuLeuPheSerGlyAsnLysAlaThrGlyProSerAsnSerSerAlaAsnGlnGluGly 422  
QY 1147 CAGCAGGGAGTG------TCATCATCTACTTCAACA- 1176  
Db 423 CysGlyGlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysLysGlyLeu 442  
QY 1177 -----GAAGATCCAGTAGTAAGTTTTCAGAAATACTGCG- 1212  
Db 443 TrpIleGluAspAsnGluAsnValSerSerGlyAsnThrAlaThrValSerGlyGly 462  
QY 1213 -----GTAGAGTTGATGGGAAC 1230  
Db 463 AlaIleTyAlaThrLysCysAlaLeuHisGlyAsnThrThrLeuThrPheAspGlyAsn 482  
QY 1231 GTAGCCCGA---GTAGGAGGGGATTAC----- 1257  
Db 483 ThrAlaGluThrAlaGlyGlyAlaIleTyThrGluThrGluAspPheThrLeuThrGly 502  
QY 1258 TCCTACGGGAACGTTGCTTCTCGTAATAAT----- 1287  
Db 503 SerThrGlyThrValThrPheSerThrAsnThrAlaLysThrAlaGlyAlaLeuHisThr 522  
QY 1288 ---GGAAAACCTTTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAGCAA 1344  
Db 523 LysGlyAsnThrSerPheThrLysAsnLysAla-----LeuValPheSerGlyAsn 539  
QY 1345 CCAACAAGTGGACAGGCTTCTAATACGAGTAATAATTACGGAGATGGAGGAGCTATCTTC 1404  
Db 540 SerAlaThrAlaThrAlaThrThrThrAspGlnGluGlyCysGlyGlyAlaIleLeu 559  
QY 1405 TGT-----AAGAATGGTGG 1419  
Db 560 CysAsnIleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAsnGluSer 579  
QY 1420 CNAAGCAGGATCCAATAACTCTGGATCTTCC----- 1452  
Db 580 LeuSerPheIleAsnAsnThrAlaLysArgSerGlyGlyIleTyAlaProLysCys 599  
QY 1453 ---TTTGATGGAGAGGAGTAGTCTTTTCTAGCAATGTAGCTGCTGGGAAAGGGGA 1509  
Db 600 ValIleSerGlySerGluSerIleAsnPheAspGlyAsnThrAlaGluThrSerGlyGly 619  
QY 1510 GCTATTATGCCAAAAGCTCTCGTGTCTAACTGTGGCCCTGTGACAATTTTAAAGGAAT 1569  
Db 620 AlaIleTySerLysAsnLeuSerIleThrAlaAsnGlyProValSerPheThrAsnAsn 639  
QY 1570 ATCGCTAATGATGGTGGAGGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCT 1629

Db 640 SerGlyGlyLysGlyGlyAlaIleTyIleAlaAspSerGlyGluLeuSerLeuGluAla 659  
QY 1630 GATTATGAGATATTTTTCGATGGCAATCTTAAAGAACAGCAAGAGAAATGCTGCC 1689  
Db 660 IleAspGlyAspIleThrPheSerGlyAsn-----ArgAlaThrGluGlyThrSer 676  
QY 1690 GATGTTAATGGCTAACTGTCTCTCAAGCCATTTTCGATGGATCGGAGGAGGAATA 1749  
Db 677 ThrProAsn-----SerIleHisIleGlyAlaGlyAlaLysIle 689  
QY 1750 ACGACATTAAAGACTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAATG--- 1806  
Db 690 ThrLysLeuAlaAlaProGlyHisThrIleTyPheTyAspProIleThrMetGlu 709  
QY 1807 -----GCAACCGA----- 1815  
Db 710 AlaProAlaSerGlyGlyThrIleGluGluLeuValIleAsnProValValLysAlaIle 729  
QY 1816 -----AATAACCGCCAGCGCAGTCTTCCAAACTTCTTAAATAATTAAC 1857  
Db 730 ValProProGlnProLysAsnGlyProIleAlaSerValProValProValAla 749  
QY 1858 GATGTCGAAGGATACACAGGGGATATTGTTTT----- 1890  
Db 750 ProAlaAsnProAsnThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAsp 769  
QY 1891 -----GCTAATGGAAGCAGTACTTTGTACCAAAATGTTAGATAGACAAAGGA 1938  
Db 770 AlaSerIleProAlaAsnThrThrIleLeuAsnGlnLysIleAsnLeuAlaGlyGly 789  
QY 1939 AGGATTGCTTCTGTGAAGGCAAAATTTATCAGTGAATCTCTAAGTCACACAGGTGG 1998  
Db 790 AsnValValLeuLysGlyAlaThrLeuGlnValTySerPheThrGlnGlnProAsp 809  
QY 1999 AGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGAATCTCCACACCCACAA 2055  
Db 826 SerThrValPheMetAspAlaGlyThrLeuGluThrThrThrThrAsn----- 826  
QY 2056 CAGCCTCTCGCGCTAATCAGTTGATCAGCTTTTCCAAATCTGCATTTGCTCTTCTTCT 2115  
Db 827 -----AsnThrAspGlySerIleAspLeuLysAsnLeuSerValAsnLeuAspAla 843  
QY 2116 TTGTAGCAACAATGCGATTACGAATCTCTACCAATCTCTCCAGGCGCAAGATCTCAT 2175  
Db 844 LeuAspGlyLysArgMetIleThr----- 851  
QY 2176 CCTGAGTCAATTGCTAGCACAACTGCTGTTCTGTACAAATTAGTGGGCTATCTTTT 2235  
Db 852 -----IleAlaValAsnSerThrSerGlyLeuLysIleSerGlyAspLeuLysPhe 869  
QY 2236 GAGGATTTGGATGATACAGCTTATGATAGGTATGATTTGGCTAGGTTCTTAATCAAAAATC 2295  
Db 870 HisAsnAsnGluGlySerPheTyAspAsn-----ProGlyLeuLysAlaAsnLeu 886  
QY 2296 AATGCTCGAAATTAACAGTTA---GGGACTAAGCCCCCAGCTAAT----- 2337  
Db 887 AsnLeuProPheLeuAspLeuSerSerThrSerGlyThrValAsnLeuAspPheAsn 906  
QY 2338 ---GCCCATCAGATTTGACTCTAGGGAATGAGATGCTAAGTATGGCTATCAAGGAAGC 2394  
Db 907 ProIleProSerSerMetAla-----AlaProAspTyThrGlyGlnGlySer 922  
QY 2395 TGGAGCTTGGCTGGGATCTTAATACAGCAAAATATGCTCTTATCTTCTCAAGCTACA 2454  
Db 923 TrpThrLeuVal-----ProLysValGlyAlaGlyGlyLysValThrLeuValAlaGlu 940  
QY 2455 TGGACTAAACTGGGTATAATCTGGGCTCAGCGAGTAGTCTTCTTGGTTCCAAATAGT 2514  
Db 941 TrpGlnAlaLeuGlyTyThrProLysProGluLeuArgAlaThrLeuValProAsnSer 960  
QY 2515 TTATGGGATCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGAT 2574  
Db 961 LeuTrpAsnAlaTyValAsnIleHisSerIleGlnGlnIleAlaThrAlaMetSer 980

QY 2575 GGGCGCTCTATTGTGCGAGGATATGCGTTTCTGGAGTTTCGAATTTCTTCTATCATGAC 2634  
Db 981 AspaLaProSerHisProGlyIleIlePheGlyGlyIleGlyAsnAlaPheHisGlnAsp 1000  
QY 2635 CCGCATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATTCCTTAGGAGCAAC 2694  
Db 1001 LysGlnLysGluAsnAlaGlyPheArgLeuIleSerArgGlyTyxIleValGlyGlySer 1020  
QY 2695 TCTACTTTGGATCATCG---ATGTTTGTCTAGCAATTTACCGAAGTATTTGGTAGACT 2751  
Db 1021 MetThrThrProGlnGluTyxThrPheAlaValAlaPheSerGlnLeuPheGlyLysSer 1040  
QY 2752 AAGATTATGTAGTGTGCTTCCATCATCATGCTTGCATAGATCCGTTTATCTATCT 2811  
Db 1041 LysAspTyxValValSerAspIleLysSerGlnValTyxAlaGlySer-----LeuCys 1058  
QY 2812 ACCAACAAAGCTTTATGTGATCTTATTTGTCGGAGATCCGTTTATCCGTGCT----- 2865  
Db 1059 AlaGlnSerSer-TyrValIleProLeuHisSerSerLeuArgArgHisValLeuSerLys 1078  
QY 2866 -----AGCTACGG 2873  
Db 1078 sValLeuProGluLeuProGlyGluThrProLeuValLeuHisGlyGlnValSerTyxG 1098  
QY 2874 GTTTCGGAATCAGCATATGAACCTCATATACATTTGCGAGAGGAGCGATGTTGTTG 2933  
Db 1098 yArgAsnHisHisAsnMetThrThrLysLeuAlaAsnAsnThrGlnGlyLysSerAspTr 1118  
QY 2934 GATATAAATGCTGCTCGCGAGAGATTTGGAGCGGATTTACCGATTTGTGATTTACTCCATC 2993  
Db 1118 pAspSerHisSerPheAlaValGluValGlyGlySerLeuProValAspLeu-----As 1136  
QY 2994 TAAGCTTATTTGAATGAGTGTGCTCTTTCGTCGCAAGCTGATTTTCTTATCCCATCA 3053  
Db 1136 nTyxArgTyxLeuThrSerTyxSerProTyxValLysLeuGlnValValSerValAsnG 1156  
QY 3054 TGAATCTTTACAGAGAGGCGCATCAAGCTCGGCATTTCAAGAGCGGACATCTCTTAA 3113  
Db 1156 nLysGlyPheGlnGluValAlaAlaAspProArgIlePheAspAlaSerHisLeuValAs 1176  
QY 3114 TCTATCAGTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTAATAATA 3173  
Db 1176 nValSerIleProMetGlyLeuThrPheLysHisGluSerAlaLysProProSerAlaLe 1196  
QY 3174 TAGCTTTATGCGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTGTTACTGAGCAAC 3233  
Db 1196 uLeuLeuThrLeuGlyTyxAlaValAspAlaTyxArgAspHisProHisCysLeuThrSe 1216  
QY 3234 GCTCCTATCCATCAGAGACATGACACAGATGCTTTCAATTTAGCAGACATGGAGT 3293  
Db 1216 r---LeuThrAsnGlyThrSerTrpSerThrPheAlaThrAsnLeuSerArgGlnAlaPh 1235  
QY 3294 TGTGTTAGAGGATCTATGATGCTTCTCTCAACAAAGTAATAGATATATAGCCATGG 3353  
Db 1235 ePheAlaGluAlaSerGlyHisLeuLysLeuLeuHisGlyLeuAspCysPheAlaSerG 1255  
QY 3354 AAGATATGAGTATGAGATGCTTCTCGAGGCTATGTTGATGTCGACGAAGTATGAGTCG 3413  
Db 1255 ySerCysGluLeuArgSerSerArgSerTyxAsnAlaAsnCysGlyThrArgTyxSe 1275  
QY 3414 GTTC 3417  
Db 1275 rPhe 1276

## RESULT 23

B72078

Polymorphic outer membrane protein g/i family - Chlamydothila pneumoniae (strain CWL029)  
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
C;Date: 23-Apr-1999 #sequence revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: B72078  
R;Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999

A;Title: Comparative Genomes of Clamydia pneumoniae and C. trachomatis.

A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: B72078

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-1407 &lt;ARN&gt;

A;Cross-references: UNIPROT:Q92899; UNIPARC:UPI0000047086; GB:AE001363; NIT

A;Experimental source: strain CWL029

C;Genetics:

A;Gene: pmp\_6

Alignment Scores: 3.5e-60 Length: 1407  
Pred. No.: 334  
Score: 991.50 Matches: 334  
Percent Similarity: 40.3% Conservative: 150  
Best Local Similarity: 27.8% Mismatches: 440  
Indels: 277  
Query Match: 2  
DB: 40

US-10-701-844-1 (1-4435) x B72078 (1-1407)

QY 376 TGTGTGATGCAACGCTCTTCCATAGTCTCTTTCTTTCTTCA-----ATG 417  
Db 299 CysTyxLeuAlaThrAlaThrAspLysThrGlyLeuAlaIleSerGlnAsnGlnGluMet 318  
QY 418 ATTCTAGCTTATCTTGTCTCTTTAAATGGGGGGA---TATGCAGCAGAAATCATG 474  
Db 319 SerPheThrSerAsnThrThrThrAlaAsnGlyGlyAlaIleTyxAlaThrLysCysThr 338  
QY 475 GTTCTCTCAAGGAATTTACGATGGGAGAGCTTAATCTGTATCATTTCCCTATATCTGTATA 534  
Db 339 Leu-----AspGlyAsnThr---ThrLeuThrPheAspGlnAsnThrAla 352  
QY 535 GGAGATCCGAGTGGGACTACTCTTTTCTGCGAGGAGCTTAACATTAATAAATCTTGAC 594  
Db 353 ThrAlaGlyCysGlyAlaIleTyxThrGluThrGlu----- 365  
QY 595 AATCTATTGTCAGCTTTCCTTAAAGTCTTTTGGGAATTTATAGGAGTCTTACTGTT 654  
Db 366 -----AspPheSerLeu 369

QY 655 TTAGGAGAGGACATCTGTTGATCTTTCGAGACATACGAGCTTCTCAAAATGGGAGCT 714  
Db 370 LysGlySerThrGlyThrValThrPheSerThrAsnThrAlaLysThrGlyGlyAlaLeu 389  
QY 715 CTAAGT-----AATAGCGCTCTGATGCA-----CTGTTTACTATTGAGGCT 756  
Db 390 TyxSerLysGlyAsnSerSerLeuThrGlyAsnThrAsnLeuLeuPheSer-----Gly 407  
QY 757 TTTAAAGAATTTATCTCTTTTCCAAATTCGAATTCATTACTTGGCGTACTGCTGCTGCAACG 816  
Db 408 AsnLysAlaThrGlyProSerAsn-----SerSer 417

QY 817 ACTAATAAGGTAGGAGCTCCGAGACACATCTACACCGCTTAATGCTACTATTTAT 876  
Db 418 AlaAsnGlnGluGlyCysGlyGlyAlaIleLeuAlaPheIleAspSerGlySerValSer 437  
QY 877 TCTAAACAGATCTTTGTTTACTCAATAATGAGAGTCTCTCATCTATATATTTAGTC 936  
Db 438 AspLysThrGlyLeuSerIleAlaAsnAsnGlnGluValSerLeuThrSerAsnAlaAla 457  
QY 937 TCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGCTTCAAGCAATTTACAGCTT 996  
Db 458 ThrValSerGlyGlyAlaIleTyxAlaThrLysCysThrLeuThrGlyAsnGlySerLeu 477

QY 997 TGTGTCTTCCAGAAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACC--- 1053  
Db 478 ---ThrPheAspGlyAsnThrAlaGlyThrSerGlyGlyAlaIleTyxThrGluThrGlu 496  
QY 1054 AGTTCTCTGCTATGCTTAACAGAGGCTCTTATGCTCTTTGCTAGCAATGTTGAGGAGTA 1113  
Db 497 AspPheThrLeuThrGlySerThrGlyThrValThrPheSerThrAsnThrAlaLysThr 516

```
QY 1114 AGAGGGGAGGAGT----- 1128
Db      ||||| :
517 ---GlyGlyAlaLeuTyrSerLysGlyAenAenSerLeuSerGlyAenThrAenLeuLeu 535
1129 -----GCTGCTGTTCCAGGATGGGAGCAG 1152
QY      ||||| :
Db      ||||| :
536 PheSerGlyAenLysAlaThrGlyProSerAenSerSerAlaAenGlnGluGlyCysGly 555
1153 CGAGTG-----TCATCATCTACTCAACA----- 1176
QY      ||||| :
Db      ||||| :
556 GlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysLysGlyLeuTrpIle 575
1177 ---GAGATCCAGTAGTAGTTTCCAGAAATCTGCG----- 1212
QY      ||||| :
Db      ||||| :
576 GluAenAenGluAenValSerLeuSerGlyAenThrAlaThrValSerGlyAlaIle 595
1213 -----GTAGAGTTTGTATGGGAACGTAGCC 1236
QY      ||||| :
Db      ||||| :
596 TyrAlaThrLysCysAlaLeuHisGlyAenThrThrLeuThrPheAspGlyAenThrAla 615
1237 CGA---GTAGGAGGAGGATTTAC-----TCCTAC 1263
QY      ||||| :
Db      ||||| :
616 GluThrAlaGlyGlyAlaIleTyrThrGluThrGluAenPheThrLeuThrGlySerThr 635
1264 GGAACGTTGCTTCTCGAATAAT-----CGA 1290
QY      ||||| :
Db      ||||| :
636 GlyThrValThrPheSerThrAenThrAlaLysThrAlaGlyAlaLeuHisThrLysGly 655
1291 AAAACCTTGTCTCAACATGTTCTCTCTGTTTACATGCTGCTAAGCAACCAACA 1350
QY      ||||| :
Db      ||||| :
656 AenThrSerPheThrLysAenLysAla-----LeuValPheSerGlyAenSerAla 672
1351 AGTGACAGGCTTCTAATACGAGTAATAATACGAGATGAGGAGCTATCTCTGT--- 1407
QY      ||||| :
Db      ||||| :
673 ThrAlaThrAlaThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 692
1408 -----AAGAATGGTGGCGCAAGCA 1425
QY      ||||| :
Db      ||||| :
693 IleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAenGluSerLeuSer 712
1426 GGAATCAATAACTCTGGATCAGTTTCC-----TTT 1455
QY      ||||| :
Db      ||||| :
713 PheIleAenAenThrAlaLysAenSerGlyGlyGlyIleTyrAlaProLysCysValIle 732
1456 GATGAGAGGAGTAGTTCTTTCTTAGTAGCAATGATGCTGCTGGGAAAGGGAGCTATT 1515
QY      ||||| :
Db      ||||| :
733 SerGlySerGluSerIleAenPheAenPheAenThrAlaGluThrSerGlyAlaIle 752
1516 TATGCCAAAAGCTCTCGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCT 1575
QY      ||||| :
Db      ||||| :
753 TyrSerLysAenLeuSerIleThrAlaAenGlyProValSerPheThrAenAenSerGly 772
1576 AATGATGTGGAGCAATTATTATAGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTAT 1635
QY      ||||| :
Db      ||||| :
773 GlyLysGlyGlyAlaIleTyrIleAlaAenSerGlyGluLeuSerLeuGluAlaIleAen 792
1636 CGAGATATTATTTCATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCGCATGTT 1695
QY      ||||| :
Db      ||||| :
793 GlyAspIleThrPheSerGlyAen-----ArgAlaThrGluGlyThrSerThrPro 809
1696 AATGCGGTAACTGTCTCTCACAGCCATTTCGATGGATCGGAGGAGGAATAACGACA 1755
QY      ||||| :
Db      ||||| :
810 Aen-----SerIleHisLeuGlyAlaGlyAlaLysIleThrLys 822
1756 TTAAGAGCTAAAGCAGGATCAGATTCTCTTAAATGATCCCATCGAGATG----- 1806
QY      ||||| :
Db      ||||| :
823 LeuAlaAlaProGlyHisThrIleTyrPheTyrAenPheProIleThrMetGluAlaPro 842
1807 GCAAAACGGA----- 1815
QY      ||||| :
Db      ||||| :
843 AlaSerGlyGlyThrIleGluGluLeuValIleAenProValValLysAlaIleValPro 862
1816 -----AATAACCGCAGCGGCTCTTCCAACTTCTAAATAAATTAACGATGTT 1863
```

```
Db      ||||| :
863 ProProGlnProLysAenGlyProIleAlaSerValProValValProValAlaProAla 882
1864 GAAGGATACACAGGGGATATTGTTTTT----- 1890
QY      ||||| :
Db      ||||| :
883 AenProAenThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAenAlaSer 902
1891 -----GCTAATGGAACGAGTACTTTGTACCAAAATGTTAGCATAGAGCAAGGAGATT 1944
QY      ||||| :
Db      ||||| :
903 IleProAlaAenThrThrThrIleLeuAenGlnLysIleAenLeuAlaGlyGlyAenVal 922
1945 GTTCTTCGTGAAAAGGCAAAATATATCAGTGAATTTCTTAAGTCAGACAGGTGGCAGT--- 2001
QY      ||||| :
Db      ||||| :
923 ValLeuLysGluGlyAlaThrLeuGlnValTyrSerPheThrGlnGlnProAenSerThr 942
2002 CTGTATATGGAAGCTGGAGGTACATGGGATTTTGTAACTCCACCAACCAACAGCCT 2061
QY      ||||| :
Db      ||||| :
943 ValPheMetAspAlaGlyThrThrLeuGluThrThrThrAen----- 957
2062 CCTGCCCTAATCAGTTGATCAGCTTCCATCTGCATTTGTCTCTTCTTCTTTGTTA 2121
QY      ||||| :
Db      ||||| :
958 ---AenThrAspGlySerIleAspLeuLysAenLeuSerValAenLeuAenAlaLeuAen 976
2122 GCAACAATGCAGTTACGAATCTCTCCTACCAATCTCCAGCGCAAGATTCTCATCTGCA 2181
QY      ||||| :
Db      ||||| :
977 GlyLysArgMetIleThr----- 982
2182 GTCAATGTTAGTACCAACTGCTGTTCTGTTACAAATAGTGGGCTATCTTTTTCAGAT 2241
QY      ||||| :
Db      ||||| :
983 IleAlaValAenSerThrSerGlyLeuLysIleSerGlyAspLeuLysPheHisAen 1002
2242 TTGATCATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2301
QY      ||||| :
Db      ||||| :
1003 AenGluGlySerPheTyrAspAen-----ProGlyLeuLysAlaAenLeuAenLeu 1019
2302 CTGAAATTTACAGTTA---GGGACTAAGCCCCCAGCTAAT-----GCC 2340
QY      ||||| :
Db      ||||| :
1020 ProPheLeuAenLeuSerSerThrSerGlyThrValAenLeuAenAenPheAenProIle 1039
2341 CCATCAGATTTGACTCTAGGGAATGAGATGCTTAAGTATGCTATCAAGGAGCTGGAAG 2400
QY      ||||| :
Db      ||||| :
1040 ProSerSerMetAla-----AlaProAspTyrGlyTyrGlnGlySerTrpThr 1055
2401 CTTGCGTGGGATCTCTAATACAGCAAAATAGTCTTATCTCTGAAAGCTACATGGACT 2460
QY      ||||| :
Db      ||||| :
1056 LeuVal-----ProLysValGlyAlaGlyLysValThrLeuValAlaGluTrpGln 1073
2461 AAAACTGGTATAATCTGGGCTGAGCGAGTACTCTTTGGTTTCCAAATAGTTTATGG 2520
QY      ||||| :
Db      ||||| :
1074 AlaLeuGlyTyrThrProLysProGluLeuArgAlaThrLeuValProAenSerLeuTrp 1093
2521 GGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGCGC 2580
QY      ||||| :
Db      ||||| :
1094 AenAlaTyrValAenIleHisSerIleGlnGlnGlnIleAlaThrAlaMetSerAspAla 1113
2581 TCTTATTTGTCAGATTTATGGGTTTCTGAGTTTTCGAATTTCTTCTATCATGACCGCAT 2640
QY      ||||| :
Db      ||||| :
1114 ProSerHisProGlyIleTrpIleGlyIleGlyAenAlaPheHisGlnAenAspLysGln 1133
2641 GCTTTAGTCCAGGATATCGGTATATAGTGGGTTTATCTCTTAGGAGCAAACTCTCTAC 2700
QY      ||||| :
Db      ||||| :
1134 LysGluAenAlaGlyPheArgLeuIleSerArgGlyTyrIleValGlyLysSerMetThr 1153
2701 TTTGGATCATCG---ATGTTTGGCTAGCATTTTACCGAAGTATTTGGTAGATCTAAAGAT 2757
QY      ||||| :
Db      ||||| :
1154 ThrProGlnGluTyrThrPheAlaValAlaPheSerGlnLeuPheGlyLysSerLysAsp 1173
2758 TATGATGTTGCTGTTTCCATCATCATGCTGTCATAGGATCCGTTTATCTATCTACCCAA 2817
QY      ||||| :
Db      ||||| :
1174 TyrValValSerAspIleLysSerGlnValTyrAlaGlySer-----LeuCysAlaGln 1191
2818 CAAGCTTTATCTGATCTCTATTTGTTGGAGATCGTTTATCTCTCGTGTCT----- 2865
QY      ||||| :
Db      ||||| :
2865
```

Db	1192	SerSer-TyrValIleProLeuHisSerSerLeuArgHisValLeuSerLysValle	1211
Qy	2866	-----AGTACGGGTTGG	2879
Db	1211	uProGluLeuProGlyGluThrProLeuValLeuHisGlyGlnValSerTyrGlyArgAs	1231
Qy	2880	GAATCAGCATATGAAACCTCATATACATTTTCAGAGGAGAGCGATGTTCTGGGATAA	2939
Db	1231	nhHisAsnMetThrThrLysLeuAlaAsnAsnThrGlnGlyLysSerAspTrpAspSe	1251
Qy	2940	TAACTCTCTGGCTGGAGAGATTGGAGCGGATTTACCGATTGTGATTACTCAATCTAAGCT	2999
Db	1251	rHisSerPheAlaValGluValGlyGlySerLeuProValAspLeu-----AsnTyrAr	1269
Qy	3000	CTATTTGAATGAGTTCGCTCTTTCGTCAGCTGAGTTTCTTATTCGCCGATCATGAATC	3059
Db	1269	gTyrLeuThrSerTyrSerProTyrValLysLeuGlnValSerValAsnGlnLysGI	1289
Qy	3060	TTTTACAGAGGAGCGATCAAGCTCGGCATTTCAAGAGCGGACATCTCTTAATCTATC	3119
Db	1289	yPheGlnGluValAlaAlaAspProArgIlePheAspAlaSerHisLeuValAsnValSe	1309
Qy	3120	AGTTCCTGTTGGAGTGAAGTTTGATCGATTTCTAGTACACATCTCAATAAATATAGCTT	3179
Db	1309	rIleProMetGlyLeuThrPheLysHisGluSerAlaLysProProSerAlaLeuLeuLe	1329
Qy	3180	TATGGCGCTTATCTGTGATGCTTATCCACCATCTCTGGTACTGAGACACGCTCCT	3239
Db	1329	uThrLeuGlyTyrAlaValAspAlaTyrArgAspHisProHisCysLeuThrSer---Le	1348
Qy	3240	ATCCCATCAGACATCGGACACAGATGCTTTCATTTAGCAGACATCGGATGTGGT	3299
Db	1348	uThrAsnGlyThrSerTrpSerThrPheAlaThrAsnLeuSerArgGlnAlaPhePheAl	1368
Qy	3300	TAGAGGATCTATGTATGCTTCTTAACAAGTATATAGAAATATATGCGCATCGAAGATA	3359
Db	1368	agLualAsrGlyHisLeuLysLeuLeuHisGlyLeuAspCysPheAlaSerGlySerCy	1388
Qy	3360	TGAGTATCGAGATGCTTCTCGAGCTATGTTTGGATGCGAGGATGATCGGTTTC	3417
Db	1388	sGluLeuArgSerSerSerArgSerTyrAsnAlaAsnCysGlyThrArgTyrSerPhe	1407
RESULT 24			
H86492			
Pmp_3 [imported] - Chlamydomydia pneumoniae (strain J138)			
C;Species: Chlamydomydia pneumoniae, Chlamydia pneumoniae			
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004			
C;Accession: H86492			
R;Shirai, M.; Hirakawa, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie			
Nucleic Acids Res. 28, 2311-2314, 2000			
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.			
A;Reference number: A86491; MUID:20330349; PMID:108711362			
A;Accession: H86492			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-772 <STO>			
A;Cross-references: UNIPROT:Q9RB71; UNIPARC:UPI000000CCC33; GB:BA000008; NID:98978389; PI			
A;Experimental source: strain J138			
C;Genetics:			
A;Gene: pmp_3_2			
Alignment Scores:			
Pred. No.:	2.05e-59	Length:	772
Score:	979.50	Matches:	261
Percent Similarity:	46.2%	Conservative:	144
Best Local Similarity:	29.8%	Mismatches:	323
Query Match:	12.4%	Indels:	149
DB:	2	Gaps:	24
US-10-701-844-1 (1-4435) x H86492 (1-772)			
Qy	919	TTCTATAGTAAATTTAGTCTCTGGAGATCGGGAGCTATAGATGCTTAAGAGCTTAACGGTT	978

Db	3	PheSerLysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaLysThrLeuSerLeu	22
Qy	979	CAAGGAATTAGCAACTTTGTCTCTTCAAGAAAATACCTGCTCAAGCTGATGGGGAGCT	1038
Db	23	ThrGlyThrThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGlyGlyAla	42
Qy	1039	TGCTAAGTAGTCACAGCTTCTCTCTATGCGCTAACAGGCTCTTATTCCTTGTAGCG	1098
Db	43	IleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyGluValSerPhe-----	60
Qy	1099	AATGTTGCAGGAGTAAGAGGGGAGGATTCGTCTGCTCAGGATGGCAGCAGGAGTNG	1158
Db	60	-----	60
Qy	1159	TCATCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTTCCAGAAATACCTCGGTAG	1218
Db	61	SerAspAsnThrSerSerAspSer-----	68
Qy	1219	TTTGATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTTC	1278
Db	69	-----GlyAlaAlaIlePheThrGluAlaSerValThrIle	80
Qy	1279	CTGATAATGGAACCACTTGTCTCTCAACATGTTGCTCTCTCTCTTACATGCTGCT	1338
Db	81	SerAsnAsnAlaLysValSerPheIleAspAsn-----	91
Qy	1339	NAGCAACCAACAGTGGCAGAGCTTCTAATACAGTAATAAATACGAGATGGAGAGCT	1398
Db	92	-----LysValThrGlyAlaSerSerThrThrGlyAspMet---SerGlyGlyAla	108
Qy	1399	ATCTTCTGTAGAATGTCGCAAGCAGGATCCAAATACCTCTGGATCAGTTCTCTTGAAT	1458
Db	109	IleCysAlaTyrLys-----ThrSerThrAspThrLysValThrLeuThr	123
Qy	1459	GGAGAGGAGTAGTCTTTCTTAGTAGCAATAGTAGCTGCGGGAAGGGGAGCTATTTAT	1518
Db	124	GlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrThrAlaGlyGlyAlaIleTyr	143
Qy	1519	GCCAAAGCTCTCGGTCTGCTAATGTCGCCCTGTACAAATTTTAAAGGAATATCGCTAAT	1578
Db	144	ValLysLeuLeuLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsnSerValAsn	163
Qy	1579	-----CATGCTGGAGCGATTTATTAGGAGAACTCTCGAGAGCTCAGTTTA	1623
Db	164	GlyGlyThrAlaProLysGlyGlyAlaIleAlaGluAspSerGlyGluLeuSerLeu	183
Qy	1624	TCTGCTGATATGAGATATTTATTTTCGATGGGAATCTTAAAGAACACAGCAAGAGAT	1683
Db	184	SerAlaAspSerGlyAspIleValPheLeuGlyAsn-----ThrValThrSerThr	200
Qy	1684	GCTGCCGATGTTAATGCGCTAATGCTCTCAAGCCATTCGATGGGATCGGAGGG	1743
Db	201	ThrProGlyThrAsn-----ArgSerSerIleAspLeuGlyThrSerAla	215
Qy	1744	AAATAACACATTAAGAGCTTAAAGCAGGCGCATCATCTCTTTAATGATCCATCGAG	1803
Db	216	LysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleTyrPheTyrAspProIleThr	235
Qy	1804	ATGCCAAACGGAATAACACGAGCCGAGCTCTTCCAACTTCTTAAATAATACGATGCT	1863
Db	236	ThrGlySerSerThrThr-----ValThrAspValLeuLysValAsnGluThr	251
Qy	1864	GAAGGA-----TACAGGAGGATATGTTTTTCTGCT-----	1893
Db	252	ProAlaAspSerAlaLeuGlnTyrThrGlyAsnIleIlePheThrGlyGluLysLeuSer	271
Qy	1894	-----AATGGAAGCAGTACTTTGTACCAAAAATGTTACGATA	1929
Db	272	GluThrGluAlaAlaAspSerLysAsnLeuThrSerLysLeuLeuGlnProValThrLeu	291
Qy	1930	GAGCAAGGAGGATGTTCTTCGTGAAAAGCGCAAAATATCAGTGAATCTCTTAAGTCAG	1989
Db	292	SerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThrGlnAlaPheThrGln	311





QY 600 TATTGACGCTTTCCCTTTAAAGTTGTTTGGGAACCTATTAGGAGTGTCTTACTGTTTAGG 659  
DB 70 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGI 88  
QY 660 GAGAGGACACTGCTGACCTTTCCAGAAACATAGGACTTCTACA-----AATGGGGCAGC 713  
DB 88 yAsnHisGlyLeuTyPheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVa 108  
QY 714 TCTAAGTAATAGCGCTGCTGATGAGACTGTTTACTATTGAGGGTGTAAAGAAATATCTCT 773  
DB 108 lLeuCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128  
QY 774 TTTCAATTGCAATTCATTACTTCCCGTACTGCTGCTGCAACGACTAATAAGGGTAGCCA 833  
DB 128 e-----IleGI 130  
QY 834 GACTCCGACAGCAACATCTACACCGTCTAATGGTACTATTATTCTTAAACACAGACTTTT 893  
DB 130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTySerLysAsnAlaLeuMe 148  
QY 894 GTTACTCAATAATAGAGAGTTCTCATTTCTATAGTAATTTAGTCTCTGGAGATGGGGAGC 953  
DB 148 tLeuLeuAsnAsnTy-ValValArgPheGluGlnAsnGlnSerLysThrLysGlyAl 168  
QY 954 TATAGATGCTAAGAGCTTAACGTTCAAGGAATTAGCAAGCTTGTCTCTCCAGAA 1013  
DB 168 aIleSerGlyAlaAsnValThrIleValGlyAsnTyAspSerValSerPheTyGlnAs 188  
QY 1014 TACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACCAGTTCTCTGCTATGGCTAA 1073  
DB 188 nAlaAlaThrPheGlyGly-----AlaIleHis 198  
QY 1074 CGAGGCTCTTATTCCTTTGTAGCGAATGTTGAGGAGTAAGAGGGGAGGATGCTGTC 1133  
DB 198 rSerGlyProLeuGlnIleAlaValaAsnGlnAlaGluIleArg----- 212  
QY 1134 TGTTCCAGGTGGCAGCAGGAGGTGTCATCTACTTCAACAGAGATCCAGTAGTAAG 1193  
DB 212 ----- 212  
QY 1194 TTTTCCAGAAATACTGCGTAGAGTTTGATGGAGTAGCCGAGTAGGAGGAGGAT 1253  
DB 213 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 226  
QY 1254 TTACTCTCAGCGAAGCTTGTCTTCTGTAATAATGGAAACCTTGTCTCAACAATGT 1313  
DB 226 uTySerAspGlyAspIleAspIleAspGlnAsnAlaTy-ValLeuPheArgGluAsnGI 246  
QY 1314 TGCTTCTCTGTTTACATTGCTGTAGCAACCAAGTGGACAGCTTCTAATACGAG 1373  
DB 246 uAlaLeuThrThrAlaIle----- 252  
QY 1374 TAATAATTACGGAGATGGAGAGCTATCTTCTGTAAGAATGTTGCGCAAGCAGGATCCAA 1433  
DB 253 -----GlyLysGlyGlyAlaValCysCys-----LeuPr 262  
QY 1434 TAACTCTGGATCA-----GTTTCTCTTTGATGGAGAGGAGTAGTTTT 1475  
DB 262 oThrSerGlySerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 282  
QY 1476 CTTTAGTAGCAATAGCTGCTGGGAAAGGGGAGCTATTATGCCAAAAGCTCTCGT 1535  
DB 282 lPheGluArgAsnHisSerIleMetGlyGlyAlaIleTyAlaArgLysLeuSerII 302  
QY 1536 TGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCT-----AATGA 1580  
DB 302 eSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTyAlaAsnSerGlnAsnLe 322  
QY 1581 TGTGGAGCGGATTTATTAGGAGAACTCTGGAGAGCTCATTTATCTGCTGATATGAGGA 1640  
DB 322 uGlyGlyAlaIleAlaIleAspThrGlyGlyGluIleSerLeuSerAlaGluLysGlyTh 342

QY 1641 TATTATTTTCGATGGGAATCTTAAAGAAACAGCCAAAGAGAATGCTGCCGATGTTAATGG 1700  
DB 342 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGI 357  
QY 1701 CGTAACTGTCCTCTCACCAAGCCATTTCCATGGGATCGGAGGGAATAACGACATTAAG 1760  
DB 357 yIleHisLeuLeuGlnAsnAla-----LysPheLeuLysLeuGI 370  
QY 1761 AGCTAAAGCAGGCAATCAGATTTCTCTTAATGATCCCATCGAGATGGCAACGGAATAA 1820  
DB 370 nAlaArgAsnGlyTySerIleGluPheTyAspProIle-----Th 384  
QY 1821 CCAGCCAGCGCAGTCTTCCAAACTTCTAAATAATTAAC-----GATGGTGAAGGATA 1871  
DB 384 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404  
QY 1872 CACAGGGATATGTTTTTGTCT-----AATCGAAG 1901  
DB 404 rThrGlyThrIleLeuPheSerGlyLysSerLeuAlaAsnAspProArgAspPheLy 424  
QY 1902 CAGTACTTTGTACCAAAATGTTTACGATAGACAGCAAGAGGATTTCTTCTCGTAAAGGC 1961  
DB 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyLeuValIleLysGluGlyAl 444  
QY 1962 AAAATTATCAGTGAATTTCTTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGG 2018  
DB 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuGI 464  
QY 2019 GAGTACATGGGATTTGTAACTCCACACCAACCAACAGCCTCTCGCCGCTAATCAGTT 2078  
DB 464 yThr-----LysLeuIleAlaSerLysGluAs 473  
QY 2079 GATCAGCTTCCAAATCTGCATTTGTCTCTTCTTCTTTGTAGCAACAACATGCACTTAC 2138  
DB 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 486  
QY 2139 GAATCTCTCTACCAATCTCCAGCGCAGAGATTTCTCATCTCGCAGTCATTTGGTAGCAAC 2198  
DB 487 -----SerSerSerThrAlaAlaValIleLysAlaAsnTh 499  
QY 2199 TGCTGCT-----TCTGTTACAAATTAGTGGGCTATCTTT----- 2232  
DB 499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519  
QY 2233 -TTTGAGGATTTGGATGATACAGCTTATGATAGGTATGATTCAGTGGTCTAATCAAAA 2291  
DB 519 aTyGluAspLeuArg-----MetArgAsnSerGlnTh 530  
QY 2292 AATCAATGTCCTGAAATTAACAGTTAGGGACTAAGCCCGCAGCTAATGCCCATCAGATT 2351  
DB 530 rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlySerValThrVa 546  
QY 2352 GACTCTAGGGAATGAGATG-----CCTAAGTATGGCTATCAAGGAAGCTGGAAGCT 2402  
DB 546 lThrAlaGlyAspPheLeuProValSerProHisTyGlyPheGlnGlyAsnTrpLysLe 566  
QY 2403 TGGCTGGGATCTTAATACAGCAATAATGCTCTTATCTCTGAAAGCTACATGAGCTAA 2462  
DB 566 uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-----PheTrpAspLy 582  
QY 2463 AACTGGGTATAATCTCGGCGCTGAGCGAGTAGTCTCTTTGGTCCAAATAGTTTATGGGG 2522  
DB 582 sIleAsnTyLysProArgProGluLysGluGlyAsnLeuValProAsnIleLeuTrpGI 602  
QY 2523 ATCCATTTAGATATACGATCTGCGCATTCAGCAATTCAGCAAGGTGTGATGGCGCTC 2582  
DB 602 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuGI 622  
QY 2583 TTATTGTCGAGGATTTATGGTTCAGATTTCTCGATTTCTTCTATCATCACCCTGATGC 2642  
DB 622 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPhePheHisValSerAlaSerGI 642  
QY 2643 TTTAGTTCAGGGATATCGGTATATTAGTGGGGTATTCTCTTAGGAGCAAACTCCTACTT 2702

Db 642 uAspAsnIleArgTyrAraGHisAsnSerGlyGlyTyrValLeuSerValAsnAsnGluI 662  
 QY 2703 TGGATCATCGATGTTT--GGTCTAGCATTTTACCGAAGTATTGTGATAGTCTAAAGATTA 2759  
 Db 662 eThrProIysHisTyrThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 682  
 QY 2760 TGTAGTGTGCTGTTCCATCATCATGCTTGGCATAGGATCCGTTTATCTATCTACCCACA 2819  
 Db 682 rAlaValSerAsnAsnGluTyrArg-----MetTyrLeu----- 693  
 QY 2820 AGCTTTATGTGGATCCTATTGTTTC-----CGAGATCGCTTTATCCG 2861  
 Db 694 -----GlySerTyrLeuTyrGlnTyrThrSerLeuGlyAsnIlePheArgTy 710  
 QY 2862 TGCTAGC----- 2868  
 Db 710 rAlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnPr 730  
 QY 2869 -----TACGGTTTGGGAATCAGCATATGAAAACCTC 2900  
 Db 730 oLeuMetIlePheHisPheLeuCysAlaTyrGlyHisAlaThrAsnAspMetLysThrAs 750  
 QY 2901 ATATACATTTGCAGAGGAGCGATGTTGCTGGGATTAATACTCTCTGGCTGGAGAGAT 2960  
 Db 750 pTyrAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCy 770  
 QY 2961 TGGAGCGGGATPACCGATGTCATCTACATCTAAGCTCTATTGGAATGAGTTCGCTCC 3020  
 Db 770 sGlyGlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr 790  
 QY 3021 TTTCGTGCAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGCGCATCA 3080  
 Db 790 oPheMetLysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheLysGluThrThrAlaAs 810  
 QY 3081 AGCTCGGCAATTCAGAGCGGACATCTCTAAATCTATCAGTCTCTGTTGGAGTGAAGTT 3140  
 Db 810 pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 830  
 QY 3141 TGATCGATGTTCTAGTACACATCTTAATAATAGCTTTATGCGGCTTATATCTGTGA 3200  
 Db 830 eGluLysLeuAlaLeuSerGlnAspValLeuTyrAspPheSerPheSerTyIleProAs 850  
 QY 3201 TGCTTATCGCACCATCTCTGGTACTGAGACAACGCTCTATCCCATCAAGACATGGAC 3260  
 Db 850 pIlePheArgLysAspProSerCysGluAlaAlaLeuValIleSerGlyAspSerTrpIle 870  
 QY 3261 AACAGATGCTTTCATTTAGCAGACATGAGTGTGGTTAGAGGATCTATGATGCTTC 3320  
 Db 870 uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyRhi 890  
 QY 3321 TCTAACAGTAATATAGATGATATAGCCATCGAAGATATGATGATCGAGATGCTTCTCG 3380  
 Db 890 sPheAsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 910  
 QY 3381 AGGCTATGTTTGGTGCAGGAGTAGAGTCCGGTTC 3417  
 Db 910 gAsnTyrAsnIleAsnCysGlySerLysPheArgPhe 922

## RESULT 26

E86491  
 polynorphic outer membrane protein G family [imported] - Chlamydomphila pneumoniae (strai  
 C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349; PMID:10871362  
 A:Accession: E86491  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-922 <STO>

A:Cross-references: UNIPROT:Q929G5; UNIPARC:UPI000004708A; GB:BA000008; NID:98978378; P1  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: pmp 1  
 C:Superfamily: Chlamydomphila pneumoniae polynorphic outer membrane protein G

Alignment Scores:  
 Pred. No.: 5.99e-59 Length: 922  
 Score: 973.00 Matches: 306  
 Percent Similarity: 41.5% Conservative: 148  
 Best Local Similarity: 28.0% Mismatches: 386  
 Query Match: 12.3% Indels: 253  
 Ds: 2 Gaps: 35

US-10-701-844-1 (1-4435) x E86491 (1-922)

QY 375 TTGTGTGATGCAACAGCTCTTCCATAAGTCTTCTTTCATAT-----GAT 419  
 Db 5 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 21  
 QY 420 TCTAGCTTATTTCTGCTCTCTTTAAATGGGGGGGATATGCAGCAGAAATCATGTGTTCC 479  
 Db 22 ThrSerLeu-SerAlaThrThrIleSer-----LeuThrPr 33  
 QY 480 TCAAGGAATTTACGATGGGAGAGCTTAACGTGTATCATTTCCCTATCTACTGTATAGGAGA 539  
 Db 33 oGluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTyrAsnVal----- 50  
 QY 540 TCCGAGTGGGACTACTGTTTTTCTGCAGGAGAGTTAATCAATTAATAAATCTTGACAAATC 599  
 Db 51 -GlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 70  
 QY 600 TATTGCACTTTGCTTTTAAAGTGTGTTTGGAACTTATTAGGAGTTTTACTGTTTATAGG 659  
 Db 70 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGl 88  
 QY 660 GAGAGGACATCGTTGACTTTTCGAGAGACATACGACATCTCTACA-----AATGGGGCAGC 713  
 Db 88 YAsnHisGlyLeuTyrPheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVa 108  
 QY 714 TCTAAGTAATAGCGTCTGATGAGCTGTTTACTTATTGAGGGTTTTAAAGAATATATCCTT 773  
 Db 108 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128  
 QY 774 TTCCAATTGCAATTCATTACTTGGCGTACTGCTGCTGCAACGACTAATAAGGGTAGCCA 833  
 Db 128 e-----IleGl 130  
 QY 834 GACTCCGACGACACATCTACACCGTCTAATGGTACTATTATTCTTAAACAGATCTTTT 893  
 Db 130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTyrSerLysAsnAlaLeuMe 148  
 QY 894 GTTACTCAATATGAGAAGTTCTCATCTTATATAATTTAGTCTCTGAGAGATGGGGGAGC 953  
 Db 148 lLeuLeuAsnAsnTyrValValArgPheGluGlnAsnGlnSerLysThrLysGlyGlyAl 168  
 QY 954 TATAGATGCTAAGAGCTTAAAGTTCAAAGGAATAGCAAGCTTTGTGTCTTCCAGAAATA 1013  
 Db 168 aIleSerGlyAlaAsnValThrIleValGlyAsnTyrAspSerValSerPheTyrGlnAs 188  
 QY 1014 TACTGCTCAAGCTGATGGGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAA 1073  
 Db 188 nAlaAlaThrPheGlyGly-----AlaIleHis 198  
 QY 1074 CGAGGCTCTTATGCTTTTGTAGCGAATGTTGCGAGAGTAAGAGGGGGGAGGATGCTGTC 1133  
 Db 198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 212  
 QY 1134 TGTTCAGGATGGGACGAGGAGGTGTTCATCTACTTCTCAACAGAAAGATCCAGTAGTAG 1193  
 Db 212 ----- 212  
 QY 1194 TTTTCCAGAAATACTGCGGTAGAGTTTGTATGGGAACGTAGCCCGAGTAGGAGGAGGAT 1253

```
Db 213 -PheAlaGlnAsnThrAlaLys-:::|||||:::|||||:::GlyGlyAlaLe 226
Qy 1254 TTACTCTACGGGAACGTGCTTCTCTGAATAATGGAAAAACCTTGTCTCAACAATGT 1313
Db 226 uTySerAspGlyAspIleAspIleAspGlnAsnAlaTyValLeuPheArgGluAsnG1 246
Qy 1314 TGTCTCTCTGTTTACATTTCGTCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAG 1373
Db 246 uAlaLeuThrThrAlaIle-|||
Qy 1374 TAATAATTACGGAGATGGAGGAGCTATCTCTGTAAGAATGGTGCACACGAGATCCAA 1433
Db 253 -GlyLysGlyGlyAlaValCysCys-|||
Qy 1434 TAACCTCGATCA-:::|||||:::GTTCTCTTGATGGAGAGGAGTAGTATTT 1475
Db 262 oThrSerGlySerSerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 282
Qy 1476 CTTTAGTAGCAATGTAGCTGCTGGGAAGGGGAGCTATTTATGCCAAAAGCTCTCGGT 1535
Db 282 lPheGluArgAsnHisSerIleMetGlyGlyGlyAlaIleTyAlaArgLysLeuSerIl 302
Qy 1536 TGTCTAAGTGGCCCTGTACAAATTTTAAAGGAATATCGCT-:::|||||:::AATGA 1580
Db 302 eSerSerGlyGlyProThrLeuPheIleAsnAlaIleSerTyAlaAsnSerGlnAsnLe 322
Qy 1581 TGTGTGAGCGAATTTATTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGAATATGAGA 1640
Db 322 uGlyGlyAlaIleAlaIleAspThrGlyGlyGlyGluIleSerLeuSerAlaGlyLysGlyTh 342
Qy 1641 TATTATTTCGATGGGAATCTTAAAGAACAGCAACAGAGAAGTCTGCCGATTTAATGG 1700
Db 342 rIleThrPheGlnGlyAsn-:::|||||:::ArgThrSerLeuProPheLeuAsnG1 357
Qy 1701 CGTAACGTGCTCTCACAAAGCCATTTTCGATGGGATCGGGAGGAGAAATAACGACATTAAG 1760
Db 357 yIleHisLeuLeuGlnAsnAla-:::|||||:::LysPheLeuLysLeuG1 370
Qy 1761 AGCTAAAGCAGGCGATCAGATTTCTTTAATGATCCCATGATGGCAACCGAAATAA 1820
Db 370 nAlaArgAsnGlyTySerIleGluPheTyAspProIle-:::|||||:::Th 384
Qy 1821 CCAGCCAGCGAGCTTCTCAAACTTCTAAATAATTAAC-:::|||||:::CATGGTGAAGATA 1871
Db 384 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404
Qy 1872 CACAGGCGATATGTTTTCGT-:::|||||:::AATGGAAG 1901
Db 404 rThrGlyThrIleLeuPheSerGlyGlyLysSerLeuAlaAsnAspProArgAspPheLy 424
Qy 1902 CAGTACTTTGTACCAAAATGTCATAGACAGAGGAGGATTTCTCTCGTGAAGAGGC 1961
Db 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyLeuValIleLysGluGlyAl 444
Qy 1962 AAAATTATCAGTGAATTTCTTAAGTCACAGAGTGGAGT---CTGTATATGGAAGCTGG 2018
Db 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuG1 464
Qy 2019 GAGTACATGGGATTTTGTAACTCCACACACCAACGAGCTCTCGCGCTAATCAGTT 2078
Db 464 yThr-:::|||||:::LysLeuIleAlaSerLysGluAs 473
Qy 2079 GATCAGCGTTTCCAATCTGATTTGTCTCTTCTTCTTTGTAGCAAAACATGCGATTAC 2138
Db 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu-:::|||||::: 486
Qy 2139 GAATCTCTTACCAATCTCCAGCGCAAGATTTCTCATCTCTCAGTCAATTTGTAGACAAC 2198
Db 487 -:::|||||:::SerSerSerThrAlaAlaValIleLysAlaAsnTh 499
Qy 2199 TGTCTGGT-----TCTGTACAAATTAGTGGGCTATCTTT-:::|||||::: 2232
```

```
499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519
2233 -TTTCAGGATTTGGATGATACAGCTTATGATAGGTATGATTTGGCTAGGTTCTTAACAAA 2291
519 aTyArgGluAspLeuArg-:::|||||:::MetArgAsnSerGlnTh 530
2292 AATCAATCTCTCGAAATTTACAGTTTAGGACTAAGCCCCAGCTAATGCCCCATCAGATTT 2351
530 rPheProLeuLeuSerLeuGlu-:::|||||:::ProGlyAlaGlyGlySerValThrVa 546
2352 GACTCTAGGGAATGAGATG-:::|||||:::CCTAAGTATGGCTATCAAGNAGCTGGAAGCT 2402
546 lThrAlaGlyAspPheLeuProValSerProHisTyArgGlyPheGlnGlyAsnTrpLys 566
2403 TCGGTGGGATCTTAATACAGCAATAATGGTCTCTTATCTCTGAAGCTACATGAGCTAA 2462
566 uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-:::|||||:::PheTrpAspLy 582
2463 AACTGGGTATATCTCGGCCCTGACGAGTACGCTTCTTGTGTTCCAAATAGTTTATGGGG 2522
582 sIleAsnTyLysProArgProGluLysGluGlyAsnLeuValProAsnIleLeuTrpG1 602
2523 ATCCATTTTAGATATACGATCTGCCATTCAGCAATTCAGCAAGTGTGGTGGGCGCTC 2582
602 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuG1 622
2583 TTATGTCGAGGATATGGTTCCTGGAGTTTCGAATTTCTTCTCATCATGACCGCATGC 2642
622 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPhePheHisValSerAlaSerG1 642
2643 TTATGTCGAGGATATCGGTATATTAGTGGGGTTTATCTTCTTAGGAGCAAACTCTTACT 2702
642 uAspAsnIleAArgTyArgHisAsnSerGlyTyValLeuSerValAsnAsnGlu11 662
2703 TGGATCATCGATGTTT---GGTCTAGCATTTACGGAAGTATTGTTGGTAGATCTTAAGATTA 2759
662 eThrProLysHisTyThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 682
2760 TGTAGTGTGCTGTTCCAAATCATCATGCTTGCATAGGATCGGTTTATCTATCTACCCAACA 2819
682 rAlaValSerAsnAsnGluTyArg-:::|||||:::MetTyLeu- 693
2820 AGCTTTATGTCGATCTTATTTGTTTC-:::|||||:::GGAGATCGCTTTATCCG 2861
694 -:::|||||:::GlySerTyLeuTyGlnTyThrThrSerLeuGlyAsnIlePheArgTy 710
2862 TGCTAGC-:::|||||::: 2868
710 rAlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnPr 730
2869 -:::|||||:::TACGGTTTGGGAATCAGCATATGAAAACCTC 2900
730 oLeuMetIlePheHisPheLeuCysAlaTyGlyHisAlaThrAsnAspMetLysThrAs 750
2901 ATATACATTTCAGAGAGAGAGCGATGCTGTTCCGGATAATAACTGTCTGCTGGAGAGAT 2960
750 pTyAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCy 770
2961 TGGAGCGGGATTACCGATTGTGATTACTCATCTAAGCTCTATTGTAAGTGGTGGCTCC 3020
770 sGlyGlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr 790
3021 TTTCTGTCAGCTGAGTTTCTTATGCGGATCATGATCTTTTACGAGGAGGAGCGCATCA 3080
790 oPheMetLysLeuGlnLeuValTyAlaTyGlnGlyAspPheLysGluThrThrAlaAs 810
3081 AGCTCGGGCATTCAGAGCGGACATCTCTTAAATCTATCAGTTCTCTGCTGGAGTGAAGTT 3140
810 pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 830
3141 TGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGGCTTATATCTGTGA 3200
830 eGluLysLeuAlaLeuSerGlnAspValLeuTyArgPheSerPheSerTyIleProAs 850
```



Db 370 nAlaAspGlnGlyTyrSerIleGluPheTyrAspProIle-----Th 384  
Qy 1821 CCGCCAGCGCAGTCTTCCAACTTCTAAATAATAC-----GATGGTGAAGGATA 1871  
Db 384 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404  
Qy 1872 CACAGGGGATATGTTTGTCT-----AATGGAAG 1901  
Db 404 rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAsnAspProArgAspPheLy 424  
Qy 1902 CAGTACTTTGTACCAAAATGTTACGATAGACGAGGAGGATGTTCTTCGTGAAAGGC 1961  
Db 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyrLeuValIleLysGluGlyAl 444  
Qy 1962 AAATATTCAGTGAATCTTAAGTCAGACAGTGGAGT---CTGTATATGAAGCTGG 2018  
Db 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuGl 464  
Qy 2019 GAGTACATGGGATTTGTAACTCCACCAACACACAGCCCTCTGCCGCTAATCAGTT 2078  
Db 464 yThr-----LysLeuIleAlaSerLysGluAs 473  
Qy 2079 GATCAGCGTTTCCAATTCGATTTGTCTCTTTCTTTCTTTGTTAGCAAAATGCGATTAC 2138  
Db 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 486  
Qy 2139 GAATCTCTTACCAATCTCCAGCGCAGAGATTCTCATCTCGCAGTCATTGGTAGCACAA 2198  
Db 487 -----SerSerSerSerThrAlaAlaValIleLysAlaAsnTh 499  
Qy 2199 TGCTGGT-----TCTGTCAATAGTGGCCCTATCTTT----- 2232  
Db 499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519  
Qy 2233 -TTTGAGGATTTGCGATGATACAGCTTATGATAGTATGATGCTAGGTTCTTAATCAAAA 2291  
Db 519 aTyrGluAspLeuArg-----MetArgAsnSerGlnTh 530  
Qy 2292 AATCAATGCTCTGAATACAGTTAGGAGTAAAGCCCGAGTAAATGCCCATCAGATTT 2351  
Db 530 rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlyGlySerValThrVa 546  
Qy 2352 GACTCTAGGGAATGAGATG-----CCTAAGTATGGCTATCAAGGAAGCTGGAAGCT 2402  
Db 546 lThrAlaGlyAspPheLeuProValSerProHisTyrGlyPheGlnGlyAsnIlePylsLe 566  
Qy 2403 TGGCTGGGATCCTAAATACAGCAATATGCTTCTTATCTCTGAAAGCTACATGGACTAA 2462  
Db 566 uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-----PheTrpAspLy 582  
Qy 2463 AACTGGGTATATCTCGGGCTGAGCGAGTAGCTTCTTTGTTCCAAATAGTTATTTGGGG 2522  
Db 582 sIleAsnTyrLysProArgProGluLysGluGlyAsnLeuValProAsnIleLeuTrpGl 602  
Qy 2523 ATCCATTTTATATACATGCTGGCATTGAGCAATTCAGCAATTCAGCAAGTGTGGTGGCGCTC 2582  
Db 602 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuGl 622  
Qy 2583 TTATTGTCGAGGATTTATGGGTTTCTGGAGTTTCCAAATTTCTTCATCATGACCGCGATGC 2642  
Db 622 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPheHisValSerAlaSerGl 642  
Qy 2643 TTTAGGTACGGGATATCGGTATATATAGTGGGGTTTATTCCTTAGGAGCAAACTCTACTT 2702  
Db 642 uAspAsnIleArgTyrArgHisAsnSerGlyGlyTyrValLeuSerValAsnAsnGluIl 662  
Qy 2703 TGGATCATCGATGTT---GCTCTAGCATTTACCAAGTATTTGGTAGATCTAAAGATTA 2759  
Db 662 eThrProLysHisTyrThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 682  
Qy 2760 TGTAGTGTGCTTCCATCATCATGCTTGCATAGGATCCGTTTATCTATCTATCCCAACA 2819  
Db 682 rAlaValSerAsnAsnGluTyrArg-----MetTyrLeu----- 693

Qy 2820 AGCTTTATGTCGATCCTATTTCTTC-----GGAGATGCGTTTATCCG 2861  
Db 694 -----GlySerTyrLeuTyrGlnTyrThrSerLeuGlyAsnIlePheArgTy 710  
Qy 2862 TCTAGC----- 2868  
Db 710 rAlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnPr 730  
Qy 2869 -----TACCGGTTTGGGAATCAGCATATGAAACCTC 2900  
Db 730 oLeuMetIlePheHisPheLeuCyAlaTyrGlyHisAlaThrAsnAspMetLysThrAs 750  
Qy 2901 ATATATCATTTGCAGAGGAGCGCATGTTCTGTTGGGATAATAACTGTCTCGCGTGGAGAT 2960  
Db 750 pTyrAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCyTrpAlaIleGluCy 770  
Qy 2961 TGGAGCGGATTCACGATTGATTCATCTCATCTAGCTCTATTGAAATGAGTTGGCTCC 3020  
Db 770 sGlyGlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr 790  
Qy 3021 TTTCTGTCAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCATCA 3080  
Db 790 oPheMetLysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheLysGluThrThrAlaAs 810  
Qy 3081 AGCTCGGCGATTCAGAGCGGACATCTCTTAATCTATCATGTTCTCTGTTGGTGAAGTT 3140  
Db 810 pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 830  
Qy 3141 TGATCGATGTTCTAGTACATCTCTTAATAATATAGCTTTATGCGCGCTTATATCTGTGA 3200  
Db 830 eGluLysLeuAlaLeu\*\*GlnAspValLeuTyrAspPheSerPheSerTyrIleProAs 850  
Qy 3201 TGTCTATCCACCATCTCTGCTACTGAGACAACGCTCTCTATCCCATCAAGACATGGAC 3260  
Db 850 pIlePheArgLysAspProSerCysGluAlaLeuValIleSerGlyAspSerTrpLe 870  
Qy 3261 AACAGATGCTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGTATGCTTC 3320  
Db 870 uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyrHi 890  
Qy 3321 TCTAACAGTAATATAGATATATGCCATGGAAGATATGATATCGAGATGCTTCTCG 3380  
Db 890 sPheAsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 910  
Qy 3381 AGCTATGTTTGCAGTGCAGGAGTAGTCCGGTTC 3417  
Db 910 gAsnTyrAsnIleAsnCysGlySerLysPheArgPhe 922

## RESULT 28

E72130

polymorphic membrane protein G family CP0761 [imported] - Chlamydotheca pneumoniae (str  
C;Species: Chlamydotheca pneumoniae, Chlamydia pneumoniae  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: E72130; G81541

R;Kalmann, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21: 385-389, 1999  
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: E72130

A;Molecule type: DNA

A;Residues: 1-841 &lt;ARN&gt;

A;Cross-references: UNIPROT:Q9Z3A1; UNIPARC:UPI00000470B5; GB:AE001586; GB:AE001363; N1

A;Experimental source: strain CWL029

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28: 1397-1406, 2000

A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A;Reference number: A81500; MUID:20150255; PMID:10684935

A;Accession: G81541

A;Molecule type: DNA

A;Residues: 1-841 &lt;REA&gt;

A;Cross-references: UNIPARC:UPI00000470B5; GB:AE002235; GB:AE002161; NID:g7189672; PIDN

A: Experimental source: strain AB39, HL cells

C: Genetics:

A: Gene: pmp\_2; CP0761

C: Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G

Alignment Scores:

Pred. No.:	1.19e-56	Length:	841
Score:	939.50	Matches:	289
Percent Similarity:	39.5%	Conservative:	134
Best Local Similarity:	27.0%	Mismatches:	358
Query Match:	11.9%	Indels:	291
DB:	2	Gaps:	30

US-10-701-844-1 (1-4435) x E72130 (1-841)

```

QY 382 ATGCAAGCTCTTCCATAAGTCTTCTTCAATGATTAGCTTATTTCTGCTCT 441
DB 1 MetLysIleProLeuArgPheLeuLeuLeuSerLeuValProThrLeuSerMetSerAen 20
QY 442 TTAATGGGGGGATATGACGAGAAATCATGGTCTCTCAAGGAATTTACGATGGGAG 501
DB 21 LeuLeuGlyAlaAlaThrThrGluGluLeuSerAlaSerAenSer---PheAspGlyThr 39
QY 502 ACGTTAACTGATCATTTCCCTATCTACTGTATAGGAGATCCGAGTGGGACTACTGTTT 561
DB 40 ThrSerThrThrSerPheSerSerLysThrSerSerAlaThrAspGlyThrAsnTyrVal 59
QY 562 TCTGAGGAGAGTAAACATTTAAATAATCTTGACAACTTCTATTGACGCTTTGCTTTAAAGT 621
DB 60 PheLysAspSerValValIleGluLeuValProLysThrGlyGluThrGlnSerThrSer 79
QY 622 TGTGTTTGGGAAC-----TTATTAGGAGTCTTACTGTTTATAGGAGAGGACACTCGTTG 675
DB 80 CysPheLysAsnAspAlaAlaAlaGlyAspLeuAenPheLeuGlyGlyPheSerPhe 99
QY 676 ACTTTCGAGAACATA---CGGACTTCTACAAATGGGAGCTCTAAGTAATAGCGCTGCT 732
DB 100 ThrPheSerAsnIleAspAlaThrThrAlaSerGlyAlaAlaIleGlySerGluAla 119
QY 733 GATGACTGTTTACTATTGAGGTTTAAAGAAATATCTTTTCCAAATTTGCAATTCATTA 792
DB 120 AenLysThrValThrLeuSerGlyPheSerAlaLeuSerPhe----- 133
QY 793 CTTGCCGTACTGCTGCTGCAACGACTAATAAGGTTAGCCAGACTCCGACGACACACTCT 852
DB 134 -----LeuLysSerProAlaSerThrVal 141
QY 853 ACACCGCTCTAATGGTACTATTATTCTTAAACAGACTCTTTGTTACTCAATAATCAGAAG 912
DB 142 ThrAenGlyLeuGlyAlaIleAenValLysGlyAenLeuSerLeuLeuAenAspAsnLys 161
QY 913 TTCTCATTTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGAT---GCTAAGAGC 969
DB 162 ValLeuIleGlnAspAenPheSerThrGlyAspGlyAlaIleAenCysAlaGlySer 181
QY 970 TTAACGTTTCAAGGAATTAGAAGCTTTGTGTTCTTCCAAAGAAATACTGCTCAAGCTGAT 1029
DB 182 LeuLysIle----- 184
QY 1030 GGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGCTTAACGAGGCTCTATTGCC 1089
DB 185 -----AlaAenAsnLysSerLeuSer 191
QY 1090 TTTGTAGCAATTTGTCAGGAGTAAGAGGGGAGGATTTGCTGCTTTCAGGATGGCAG 1149
DB 192 PheIleGlyAenSerSerSerThrArgGlyAlaIleHisThr----- 206
QY 1150 CAGGGAGTGTCTATCTACTTCAACAGAAGATCCAGTAGTAAGTTTTCAGAAATACT 1209
DB 206 ----- 206
QY 1210 GCGGTAGATTGATGGGAACGTAGCCCGTAGTAGGAGGAGTTTACTTCTACGGGAC 1269
  
```

```

DB 207 -----LysAen 208
QY 1270 GTTGCTTTCTGAATAATATGAAACCTTGTCTTCTCAACAAATGTTGCTTCTCTGTTTAC 1329
DB 209 LeuThrLeuSerSerGlyGlyGluThrLeuPheGlnGlyAenThrAla----- 224
QY 1330 ATTGCTGTAAGCAACCAAGTGGACAGGCTTCTTAATACGAGTAATAATTAACGAGAT 1389
DB 225 -----ProThr----- 226
QY 1390 GGAGGAGCTATCTTCTGTAAGAATGGTGGCAAGCAGGATCCAATAAATCTGATCAGTT 1449
DB 226 ----- 226
QY 1450 TCCTTTGATGAGAGGAGTAGTATTTCTTTAGTAGCAATGTAGCTGCGGAAAGGGGA 1509
DB 227 -----AlaAlaGlyLys----- 230
QY 1510 GCTATTTATGCCAAAAGACTCTCGTGTCTAACTGTGCGCCCTGTACAATTTTAAAGGAAT 1569
DB 230 ----- 230
QY 1570 ATCGCTAATGATGGAGCGATTTATTTAGGAGAATCTGAGAGCTCAGTTTATCTGCT 1629
DB 231 -----GlyGlyAlaIleAlaAlaAspSerGlyThrLeuSerIleSerGly 246
QY 1630 GATTATGAGATATTTTTCGATGGGAATCTTAAAGAACAGCAAGAAAGAAATGCTGCC 1689
DB 247 AspSerGlyAspIleIlePheGluGly-----AsnThrIle 258
QY 1690 GATGTTAATGCGTAACGTGTCTCTCAAGCCATTTTCGATGGATGGGAGGAAATA 1749
DB 259 GlyAlaThrGly---ThrValSerHisSerAlaIleAspLeuGlyThrSerAlaLysIle 277
QY 1750 ACACATTAAGAGCTAAAGCAGGCGATCAGATTTCTTTTATGATCCCATCGATGGCA 1809
DB 278 ThrAlaLeuAlaAlaGlnGlyHisThrIleThrPheThrAspProIleThrVal--- 296
QY 1810 AACGAAATAACAGCCAGCGCTTCCAAATCTTAAATAATTAAC----- 1857
DB 297 ThrGlySerThrSerValAlaAspAla-----LeuAenIleAenSerProAspThr 313
QY 1858 ---GATGTTGAGGATACACAGGGGATATTTGTTGCT----- 1893
DB 314 GlyAspAenLysGluThrGlyThrIleValPheSerGlyGlyLysLeuThrGluAla 333
QY 1894 -----AATGGAAGCAGTACTTTGTACCAAAATGTTTACGATAGACAA 1935
DB 334 GluAlaLysAspGlyLysAenArgThrSerLysLeuLeuGlnAenValAlaPheLysAen 353
QY 1936 GGAAGGATTTGTTCTGTAAGGAAAGCAAAATATCAGTGAATTTCTTAAGTCAGACAGT 1995
DB 354 GlyThrValValLeuLysGlyAspValValLeuSerAlaAenGlyPheSerGlnAspAla 373
QY 1996 GGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCA 2052
DB 374 AenSerLysLeuIleMetAspLeuGlyThrSerLeu----- 385
QY 2053 CAACAGCCTCTCGCGCTAATCAGTTGATCAGCTTTTCCAATCTGCAATTTGTCTTTCT 2112
DB 386 -----ValAlaAenThrGluSerIleGluLeuThrAenLeuGluIleAenIleAsp 402
QY 2113 TCTTTGTTAGCAACAATGCGATTCAGAAATCCTCTACCAATCTCCAGCGCAAGATTC 2172
DB 403 SerLeuAenGlnGlyLysIle----- 410
QY 2173 CATCTCGAGTCATTTGTTAGCACAACCTGCTGTTCTGTACAAATTAGTGGGCTATCTTT 2232
DB 411 -----LysLeuSerAlaAlaThrAlaGlnLysAspIleArgIleAspArgProVal 428
QY 2233 TTTGAGGATTTGATGATACAGCTTATGATAGG-----TAT 2268
DB 429 LeuAlaIleSerAspGluSerPheThrGlnAenGlyPheLeuAenGluAspHisSerThr 448
  
```



```
QY 2269 GATTGGCTAGGTTCTTAATCAAAAATCAATGTCTCGAATTAACAGTTAGGAGCTAAGCC 2328
Db 449 App
QY 2329 CCAGCTAAATGCCCATCAGATTGACTCTAGGGAATGAGATGCTTAAGTATGCTGCTCAA 2388
Db 461 ValIleSerAlaAspSerArgSerIleAspAlaValGlnSerPro---TyrGlyTyrGln 479
QY 2389 GGAAGCTGGAAGCTTCGGTGGAGTCTTAATACAGCAAAATAATGCTCTTATCTATCGAAA 2448
Db 480 GlyLysTrpThrIleAsnTrpSerThrAspAspLys-----Lys 492
QY 2449 GCTTACA-----TGGACTAAACTGGGTATATCTCTGGGCTGAGCGAGTAGCTTCTTTG 2502
Db 493 AlaThrValSerTrpAlaLysGlnSerPheAsnProThrAlaGlnGlnAlaProLeu 512
QY 2503 GTTCCAAATAGTTTATGGGATCCATTTAGATATACGATTCGCGCATTCAGCAATTCAA 2562
Db 513 ValProAsnLeuLeuTrpGlySerPheIleAspValArgSerPheGlnAsnPheIleGlu 532
QY 2563 GCAAGTGTGATGGGCGCTCTTAATGTGAGGATTAATGGTTCTCGAGTTTCGAATTC 2622
Db 533 LeuGlyThrGluGlyAlaProTyrGluLysArgPheTrpValAlaGlyIleSerAsnVal 552
QY 2623 TTTCTATCATGACCGCATGCTTTAGTTCAGGATATCGGTATATAGTGGGTTATTC 2682
Db 553 LeuHisArgSerGlyArgGluAsnGlnArgLysPheArgHisValSerGlyAlaVal 572
QY 2683 TTAGGAGCAACCTCTACTTT---GGATCATCGATGTTTGGTCTAGCATTTACCGAAGTA 2739
Db 573 ValGlyAlaSerThrArgMetProGlyGlyAspThrLeuSerLeuGlyPheAlaGlnLeu 592
QY 2740 TTTGGTAGATCTAAAGATTATGATGTGTGCTTCCAAATCATCATGCTTCGATAGATCC 2799
Db 593 PheAlaArgAspLysAspTyrPheMetAsnThrAsnPheAlaLysThrTyrAlaGlySer 612
QY 2800 GTTATCTATCTACCCACAGCTTTATGTGGA-----TCCTATTGTTCGAGATGCG 2853
Db 613 LeuArgLeuGlnHisAspAlaSerLeuTyrSerValValSerIleLeuLeuGlyGluGly 632
QY 2854 TTTATCCGT----- 2862
Db 633 GlyLeuArgGluLeuLeuLeuProTyrValSerLysThrLeuProCysSerPheTyrGly 652
QY 2863 ---GCTAGCTACGGTTTGGGAATCAGCATATGAAATCACTCATAT----- 2904
Db 653 GlnLeuSerTyrGlyHisThrAspHisArgMetLysThrGluSerLeuProProPro 672
QY 2905 ---ACATTCAGAGGAGAGCGATGCTGTTGGATATTAACCTGCTGGCTGGAGATTT 2961
Db 673 ProThrLeuSerThrAsp---HisThrSerTrpGlyGlyTyrValTrpAlaGlyGluLeu 691
QY 2962 GAGCGGGATTACCGATTGTGATGTTACTTCCATCTAAGCTCTATTGGAATGAGTTGCGCCT 3021
Db 692 GlyThrArgValAlaValGluAsnThrSerGlyArgGlyPhePheGlnGluTyrThrPro 711
QY 3022 TTCGTGCAAGCTGAGTTTCTTATGCGCATCATCAATCTTTTACAGAGGAAGCGATCAA 3081
Db 712 PheValLysValGlnAlaValTyrAlaArgGlnAspSerPheValGluLeuGlyAlaIle 731
QY 3082 GCTCGGCATTCAGAGCGGACATCTCTCTAAATCATCATGTTCTGTTGGAGTGAATTT 3141
Db 732 SerArgAspPheSerAspSerHisLeuTyrAsnLeuAlaIleProLeuGlyIleLysLeu 751
QY 3142 GATCGATGTTCTAGTACATCTCTAATAATATAGCTTTATGGCGGCTTATATCTGAT 3201
Db 752 GluLysArgPheAla-----GluGlnTyrTyrHisValValAlaMetTyrSerProAsp 769
QY 3202 GCTTATCGCACATCTCTGGTACTGAGACACGCTCTCTATCCCATCAAGAGATGGACA 3261
Db 770 ValCysArgSerAsnProLysCysThrThrThrLeuLeuSerAsnGlnGlySerTrpLys 789
```

```
QY 3262 ACAGATGCGCTTTCAATTAGCAAGACATCGAGTGTGGTTAGGAGATCTATGATGCTTCT 3321
Db 790 ThrLysGlySerAsnLeuAlaArgGlnAlaGlyIleValGlnAlaSerGlyPheArgSer 809
QY 3322 CTAACAAGTAATATAGAAGTATATGCCATGCAAGATATGAGTATCGAGATCTTCTCCA 3381
Db 810 LeuGlyAlaAlaAlaGluLeuPheGlyAsnPheGlyPheGluTrpArgGlySerSerArg 829
QY 3382 GCGCTATGTTTTCAGTGTGAGGAGTAGAGTCCGGTTC 3417
Db 830 SerTyrAsnValAspAlaGlySerLysIleLysPhe 841

RESULT 29
F81721
polymorphic membrane protein G family TC0267 [imported] - Chlamydia muridarum (strain N1)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81721
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: F81721
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-867 <DET>
A:Cross-references: UNIPROT:Q9PL41; UNIPARC:UPI000005786B; GB:AE002294; GB:AE002160; NI
C:Genetic:
A:Experimental source: strain Nigg (MoPn)
A:Gene: TC0267
C:Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 2,78e-46 Length: 867
Score: 789.00 Matches: 268
Percent Similarity: 40.5% Conservative: 148
Best Local Similarity: 26.1% Mismatches: 411
Query Match: 10.0% Indels: 200
DB: 31

US-10-701-844-1 (1-4435) x F81721 (1-867)
QY 433 TGTGCTCTTTAAATGGGGGATATGCAGAGAAATCATGTTCTCTCAAGAAATTTC 492
Db 9 CysCys-----LeuCysAlaThrIleLeuSerProThrAlaIleLeu 22
QY 493 GATGGGAGACGTTAACTGTATCATTTCCCTATCTATAGGAGATCGAGTGGGACT 552
Db 23 PheGlyGlnAsp---AlaLeuAspLysSerAlaLeuIleThrLysAsnProAsnSerIle 41
QY 553 ACTGTTTTTCTCGAGAGAGTTACATTAATAAACTTGACAAATCTTATTCAGCTTTG 612
Db 42 ValCysThrPheLeuGluAspCysThrMetGluAsnPheSerProAlaLeuLeuSerHis 61
QY 613 -----CCTTTAAGTCTTTTGGGAACCTTATAGGAGTTTACTGTTTAA 657
Db 62 AlaArgGlnAspAspProLeuTyrIleIleGlyAsnThr----- 74
QY 658 GGGAGAGGACACCTCGTTGACTTTTCGAGACATACGCGACTTCTACAAATGGGAGCTCTA 717
Db 75 -----HisAsnTrpPheValSerAsnLeuHisProSerThrAsnGluGluArgPhe 91
QY 718 AGTAATAGCGCTCTGATGAGCTGTTTACTATTAGGGGTTTAAAGAAATTCCTTTTCC 777
Db 92 LeuLysGluLysGlyAsp-----LeuSerIleGlnAspPheArgPheLeuSerPheThr 109
QY 778 AATTGCATTCATTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACT 837
Db 110 AspCysSerSer----- 113
QY 838 CCGACGACACATCTACACCGTCTAATGGTACTATTATTCTTAAACA---GATCTTTTG 894
```

Db 114 ---SerThrGluAspSerProSer-----lleLeuTyrHisLysAsnGlyGlnLeuPhe 130  
Qy 895 TTACTCAATAATGAGAAGTCTCTCATTTATAGTAATTTAGTCTCTGGAGATGGGGAGCT 954  
Db 131 LeuArgAsnAsnGlyAsnMetSerPheTyrArgAsnHisSerGluGlySerGlyGlyAla 150  
Qy 955 ATAGATGCTAGAGCTTTACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAAAAT 1014  
Db 151 SerSerThrAspAlaLeuPheLeuGlnHisAsnTyrLeuPheThrAsnPheGluGluAsn 170  
Qy 1015 ACTGCTCAAGCTGATGGGGAGCTGTGCAAGTAGTCACCAAGTTTCTCTGCTATGCGTAAC 1074  
Db 171 SerSerAlaLysAsnGlyGlyAlaIleGlnAlaGlnThr 183  
Qy 1075 GAGGCTCTATTGCTTTGTAGCGAATGTTGCAGGAGTAAGAGGGGAGGATTGCTGCT 1134  
Db 183 ----- 183  
Qy 1135 GTTCAGGATGGCAGCAGGAGTGCATCATCTTCAACAGAAGATCCAGTAGTAAGT 1194  
Db 184 -----LeuSer 185  
Qy 1195 TTTTCCAGAAAT---ACTGCGGTAGAGTTTGATGGGAACGTAGCCGAGTAGGAGGG 1251  
Db 186 LeuSerArgAsnValSerSerLeuSerPheSerArgAsnArgAlaAsnLeuAsnGlyGly 205  
Qy 1252 ATTTACTCTACGGAACTGCTTTCTGTAATATGGAATAACCTGTTTCTCAACAAT 1311  
Db 206 AlaIleCysCysGlnAsnLeuIleCysSerGlyAsnValAsnProLeuPhePhe----- 223  
Qy 1312 GTTGCTTCTCTGTTTACATTCGCTGCTAAGCAACCAAGTGGACAGGCTTCTAATAG 1371  
Db 224 -----Thr 224  
Qy 1372 ACTAATAATTACGGAGATGGAGAGCTATCTTCTGAAGATGGTGGCAGCAGGATCC 1431  
Db 225 AsnAsnSerAlaLeuAsnGlyGlyAlaIleCysCysIleAsnGlnGlnAsnLeuSerGlu 244  
Qy 1432 AATAACTCTGGAATCTGTTTCTGATGGAGGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
Db 245 LysGlyCysLeuSerLeuAlaTyrAsnGlnGlu-----ThrLeuPheSerGlyAsnSer 262  
Qy 1492 GCTGCTGGGAAGGGGAGCTATTTATGCCAAAAAGCTCTCGGTGTGCTAACTGTGGCCT 1551  
Db 263 AlaLysGluLysGlyGlyAlaIleTyrThrLysHisMetValLeuArgHisAsnGlyPro 282  
Qy 1552 GPACAAATTTTAAGGAATATCCTAATGATGCTGAGCGATTTATTAGGAGAATCTGA 1611  
Db 283 ValSerPheValAsnAsnSerAlaLysLeuGlyGlyAlaIleAlaIleGlnSerGlyGly 302  
Qy 1612 GAGCTCAGTTTATCTGCTGATTATGAGATATATTTTCGATGGGAATCTTAAAGAACA 1671  
Db 303 SerLeuSerIleIleAlaGlyGlySerValLeuPheGlnAsnAsnSerCysHisPhe 322  
Qy 1672 GCCAAGAGAATGCTGCCGATTTAATGGCGTAATCTGCTCTCAAGCCATTCGATG 1731  
Db 323 SerAspGlnGlyThrVal-----ArgAsnAlaIleTyrLeu 334  
Qy 1732 GATCGGGGGAATAACGACATTAAGCATTAAGCTAAAGCAGGCGATCAGATTTCTTTAAT 1791  
Db 335 GluLysAsnAlaLeuLeuSerSerLeuGluAlaArgHisGly-----AspIleLeuPhePhe 353  
Qy 1792 GATCCCATCGATGGCAACGAATAACACGACGACGACGCTTTCCTCAAACTTCTAAA 1851  
Db 354 AspProIleValGlnGluValSerProGluPheSerThrSerAlaLeuThrPro 373  
Qy 1852 ATTAACGATGGTGAAGGATACACAGGGGATATGTTTTT----- 1890  
Db 374 LeuArgIleGlnThrAsnThrAsnArgAlaValIlePhePheSerSerGluAsnLeuSerLys 393  
Qy 1891 -----GCTAATGGAGCAGTACTTTGTACCAAAAATGTTACGATAGACAA 1935  
Db 394 GluGluLysThrGluAlaAsnLeuIleSerLysIleGlnGlnProIleGluLeuGlnSer 413

Qy 1936 GGAAGGATTTGTTCTTCGTGAAAAGGCAAAATTATATCACTGAATTTCTTAAGTCACACAGGT 1995  
Db 414 GlyCysLeuValLeuLysAspArgValIleLeuSerAlaProSerLeuSerGlnAlaPro 433  
Qy 1996 GGGAGT---CTGTATATGGAAGCTGGCAGTACATGGGATTTTGTAACTCCACAACCA 2052  
Db 434 GlnAlaLeuLeuValMetAspValGlyThrSer----- 444  
Qy 2053 CAACAGCCTCTCGCGCTTAATCACTGATCAGCTTTCCCAATCTGCATTTGCTCTTTCT 2112  
Db 445 -----LeuThrThrSerSerAspLeuLysLeuThrLeu 456  
Qy 2113 TCTTTGTTAGCAACAATGCGATTAACAATCTCTCAATCTCCAGCAAGATTCT 2172  
Db 457 SerIleProLeuHisSerIle----- 463  
Qy 2173 CATCTCGAGTCATTGTTAGCACAACTGCTGTTCTGTGTACAATTAGTGGGCCT----- 2226  
Db 464 -----AspThrGluAsnSerValSerIleGlnSerProThrLeu 476  
Qy 2227 -----ATCTTTTGTAGGATTTGAT---GATACAGCTTTATGATAGGTATGAT 2271  
Db 477 SerIleGlnLysIlePheLeuSerAsnSerGluHisGluAsnPheTyrGluAsnValGlu 496  
Qy 2272 TGGCTAGTCTTAATCAAAA---ATCAATGCTCTGAAATTTACAGTTAGGACATAAGGCC 2328  
Db 497 LeuLeuSerLysAspGlnLysAspIleProLeuLeuSerLeuProLysGlyLeuProHis 516  
Qy 2329 CCAGCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCCCTAAGTAGTGGCTATCAA 2388  
Db 517 Pro-----AspLeuProAspGlyAsnLeuSerSerHisPheGlyTyrGln 531  
Qy 2389 GGAAGCTGGAAGCTTGGTGGGATCCTAATAACAGCAATAATGGTCTTATCTCTGAAA 2448  
Db 532 GlyAspTrpAsnPheSerTrpGlnThrSerAspGlnArgGlu-----ThrLeuVal 548  
Qy 2449 GCTACATGAGCTAAACCTGGGTATACTCTGGGCTGAGCAGTAGTCTTTTGGTTCCA 2508  
Db 549 AlaAsnTrpThrAlaAsnSerTyrIleProHisProGluArgGlnSerAlaLeuValAla 568  
Qy 2509 AATAGTTATCGGATCCATTTTAGATATACATCGATCTGCGCATTCAGCAATTCAGCAAGT 2568  
Db 569 AsnThrLeuTrpAsnThrTyrSerAspMetGlnAlaValGlnSerMetIleAsnThrThr 588  
Qy 2569 GTGGATGGCGCTCTTATTGTCGAGGATTATGGGTTCTCGAGTTTCGAAATTTCTTCTAT 2628  
Db 589 AlaGlnGlyGlyAlaTyrLeuPheGlyThrTrpGlySerAlaValSerAsnLeuPheTyr 608  
Qy 2629 ---CATGACCCGATGCTTTTAGT---CAGGATATCGGTATATATAGTGGGGTTATTCC 2682  
Db 609 SerHisGlyAsnSerGlyLysSerThrAspAsnTrpLysHisArgSerLeuGlyTyrLeu 628  
Qy 2683 TTAGAGCAAACTCCTAC---TTTGGATCATCGATGTTGCTGCTAGCATTTACCGAATA 2739  
Db 629 PheGlyIleSerThrHisSerLeuAspHisSerPheCysLeuAlaIleGlyGlnLeu 648  
Qy 2740 TTTGTAGATCTAAAGATTATGATGTGTCTTCCCAATCATCATCTTGATAGATGAT 2799  
Db 649 PheGlyLysSerSerAspSerPheValThrSerAlaAspThrSerTyrIleAlaAla 668  
Qy 2800 GTTTATCTATCAACCACAAAGCTTTATGTGATCCTATTGTTGTCGGAGATGCGTTATC 2859  
Db 669 Ile-----GlnThrGlnIleAlaThrSerLeuIleLysIleSerAla----- 682  
Qy 2860 GTGCTAGCTACGGTTGGGAATCAGCATATGAAACCTCATAT---ACATTTCCAGAG 2916  
Db 683 GlnAlaCysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSerPheSerLys 702  
Qy 2917 GAGACCGATGTTGTTGGGATAATAACTGTCTGCTGGAGGATTTGGAGCGGATTTACCG 2976  
Db 703 GluGlyPheGlyAlaTrpHisSerValAlaValSerGlyGluIleGlyAlaSerIlePro 722

```
QY 2977 ATTGATGATTACTCCATTAAGCTCTATTGAATGAGTTGGCTCTCTTTGCGCAAGCTGAG 3036
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
723 IIVaISerAnGlySerGlyLeuPhe---SerSerPheSerIlePheSerIleLeuGln 741
QY 3037 TTTTCTTATGCCGATCATGAATCTTTTACAGAGAAAGCGGATCAAGCTCGGGCATTCAG 3096
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
742 GlyPheSerGlyLeuGlnAspGlyPheGluGluSerArgGlyGluAlaArgAlaPheAla 761
QY 3097 AGCGGACATCTCTAATCTCATGTTCCCTGTTGGAGTGAAGTTTGATCGATGTTCTAGT 3156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
762 AspSerSerPheThrAenIleSerLeuProValGlyIleAlaPheGluGlySerGln 781
QY 3157 ACATCTCTAATAATAGCTTTATGCGCGCTTATCTGTGATGCTTATCGCACCATC 3216
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
782 LysThrArgAnTyTyHisPheLeuGlyAlaTyIleGlnAspLeuLysArgCysVal 801
QY 3217 TCTGGTACTGAGCAACGCTCTATCCATCAAGAGACATGGACACAGATGCTTTCAT 3276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
802 GluSerGlyProValThrLeuLeuLysAsnSerValThrTrpAspAlaProMetAlaAsn 821
QY 3277 TTAGCAAGACATGGAGTTGGTTAGGAGTCTATGTATGCTTCTCTAACAAGTAATATA 3336
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
822 LeuAspSerArgAlaTrpMetPheArgLeuThrAsnGlnArgAlaLeu---HisArgPhe 840
QY 3337 GAAGTATATGCCATGGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGAGT 3396
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
841 GlnThrLeuValAsnMetSerTyMetLeuArgGlyGlnSerTySerTySerLeuAsp 860
QY 3397 GCAGGAAGTAGAGTCCGGTTC 3417
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
861 LeuGlyThrThrTyArgPhe 867

RESULT 30
B71460
probable outer membrane protein 1 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: B71460
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: B71460
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-878 <ARN>
A:Cross-references: UNIPROT:O84882; UNIPARC:UPI0000131CF7; GB:AE001361; GB:AE001273; NID
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: pmpl
C:Superfamily: Chlamydophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 4.47e-45 Length: 878
Score: 771.50 Matches: 268
Percent Similarity: 40.5% Conservative: 145
Best Local Similarity: 26.5% Mismatches: 390
Query Match: 9.8% Indels: 207
DB: 2 Gaps: 35

US-10-701-844-1 (1-4435) x B71460 (1-878)
QY 538 GATCGAGTGGACTACTGTTTTTCTCGCAGGAGTTAACATTAAATAAATCTTGACAT 597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
26 AspProLeuGlyGluThrAla-----LeuLeuThrLysAsnProAsnHis 40
QY 598 TCTATTGACGCTTTCCTTTAAGTTGTTTGGAACTTATTA----- 639
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 ValValCysThrPhePheGluAspCysThrMetGluSerLeuPheProAlaLeuCysAla 60
QY 640 -----GGGAGTTTACTCTTTTAGGGAGGAGACACTCGTTGACTTTCGAG 684
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 HisAlaSerGlnAspAspProLeuTyRValLeuGlyAsnSerTyTyCysTrpPheValSer 80
```

```
QY 685 AACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTT 744
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 LysLeuHisIleThrAspProLysGluAlaLeuPheLysGluLysGlyAsp-----Leu 98
QY 745 ACTATTGAGGGTTTTAAAGAAATATTCCTTTTCCAATTGCAATTCATTACTTCCCGTACTG 804
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 SerIleGlnAsnPheArgPheLeuSerPheThrAspCysSer----- 113
QY 805 CCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACAACTCTACACGCTTAAT 864
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 -----LysGluSerSerProSer--- 119
QY 865 GGTACTATTATTCTAAACA---GATCTTTTGTACTCAATAATAGAGAGTTCTCATTC 921
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 ---IleIleHisGlnLysAsnGlyGlnLeuSerLeuArgAsnAsnGlySerMetSerPhe 138
QY 922 TATAGTAATTTAGTCTCTCGAGATGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
139 CysArgAsnHisAlaGluGlySerGlyAlaIleSerAlaAspAlaPheSerLeuGln 158
QY 982 GGAATTTAGCAGCTTTGTCTCTTCCAAGAAATACTGCTCAAGCTGATGGGAGCTTGT 1041
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 HisAsnTyLeuPheThrAlaPheGluGluAsnSerSerLysGlyAsnGlyAlaIle 178
QY 1042 CAAGTAGTCACCAAGTTTCTCTGCTATGGCTAACAGGCTCTTATTCCTTTAGCGAAT 1101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 GlnAlaGlnThr---PheSerLeuSerArgAsnValSerPheIleSerPheAlaArgAsn 197
QY 1102 GTTGCGAGGATAAGAGGGGAGGATGCTGCTGTTCCAGGATGGCAGGAGGTGCTCA 1161
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 ArgAlaAspLeuAsnGlyAlaIleCysCys----- 208
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACTCGGTAGAGTTT 1221
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
208 ----- 208
QY 1222 GATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTTACGGGAACGTTGCTTTCCTG 1281
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 -----SerAsnLeuIleCysSerGlyAsnVal----- 217
QY 1282 AATAATGGAAAAACCTTGTCTTCTCAAAATGTTGCTTCTCTGTTTACATTGCTGCTAAG 1341
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
218 -----AsnProLeuPhePhe----- 222
QY 1342 CAACCAACAAGTGGACAGCGCTTCTAATACAGTAATAATTACGAGATGAGGAGGCTATC 1401
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 -----ThrGlyAsnSerAlaThrAsnGlyGlyAlaIle 233
QY 1402 TTTCTGTAAGAATGTTGGCAAGCAGGATCCAATAACTCTCGATCAGTTTCTCTTGTATGGA 1461
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 CysCysIle-----SerAspLeuAsnThrSerGluLysGlySerLeuSerLeuAlaCys 251
QY 1462 GAGGAGTAGTCTTTTCTTTAGTACAAATGTAGTCTGGGAAGAGGGAGCTATTTATGCC 1521
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
252 AsnGlnGluThrLeuPheAlaSerAsnSerAlaLysGluLysGlyAlaIleTyAla 271
QY 1522 AAAAAGCTCTCGGTGCTTAAGTCTGCGCTGCTACAAATTTTAAAGGATATCGTAAATGAT 1581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
272 LysHisMetValLeuArgTyArgGlyProValSerPheIleAsnSerAlaLysIle 291
QY 1582 GGTGGAGCGATTATTATTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATGATGAGAT 1641
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
292 GlyGlyAlaIleAlaIleGlnSerGlySerLeuSerIleLeuAlaGlyGluGlySer 311
QY 1642 ATTATTTTCGATGGGAATCTTAAAGAACACGCCAAGAGAAATGTCGCCGATGTTAATGGC 1701
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
312 ValLeuPheGlnAsnAsnSerGlnArgThrSerAspGlnGlyLeuValArg---AsnAla 330
QY 1702 GTAATCTGTCTCACAGGCATTTCCATGGGATCGGAGGAGGAAAAATAACGACATTAAGA 1761
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
331 IleTyRLeuGluLysAspAlaIle-----LeuSerSerLeuGlu 343
```



DB:	2	Gaps:	33
US-10-701-844-1 (1-4435) x B86492 (1-712)			
QY	382	ATGCAACGCTCTTCCATAGTTCTTCTTCAATGATCTAGCTTATTTCTGCTCTCT 441	
DB	1	MetLysIleProLeuArgPheLeuLeuIleSerLeuValProThrLeuSerMetSerAen 20	
QY	442	TTAATATGGGGGGATATCGACGAGAAATCATGGTTCTCCTCAAGGAATTTACGATGGGAG 501	
DB	21	LeuLeuGlyAlaAlaThrThrGluGluLeuSerAlaSerAenSer---PheAspGlyThr 39	
QY	502	ACGTTAACTGATCATCTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTTTT 561	
DB	40	ThrSerThrThrSerPheSerSerLysThrSerSerAlaThrAspGlyThrAenTyrVal 59	
QY	562	TCTCGAGGAGTAACTATTAATAATCTTGCAATCTTATTCGACGTTTGCCTTTAAGT 621	
DB	60	PheLysAspSerValValIleGluAenValProLysThrGlyGluThrGlnSerThrSer 79	
QY	622	TGTTTTGGGAAC-----TTATTAGGAGTTTACTGTTTGTAGGAGGACACTCGTTG 675	
DB	80	CysPheLysAenAspAlaAlaIleGlyAspLeuAenPheLeuGlyGlyPheSerPhe 99	
QY	676	ACTTTCGAGACATA---CGAGCTCTACAAATGGGCGAGCTCTAAGTAAATAGCGCTGCT 732	
DB	100	ThrPheSerAenIleAspAlaThrThrAlaSerGlyAlaAlaIleGlySerGluAlaAla 119	
QY	733	GATGGAGCTGTTACTATTGAGGGTTTAAAGAATTTATCTTTTCCCAATTCGAATTCATTA 792	
DB	120	AsnLysThrValThrLeuSerGlyPheSerAlaLeuSerPhe----- 133	
QY	793	CTTGCCGTACTGCTGCTGCAACGACTAATAAGGTTAGCCAGACTCCGACGACACACTCT 852	
DB	134	-----LeuLysSerProAlaSerThrVal 141	
QY	853	ACACCGTCTAATGCTACTATTATCTTAAACAGATCTTTGTGTACTCAATAATGAGAAG 912	
DB	142	ThrAenGlyLeuGlyAlaIleAenValIleGlyAenLeuSerLeuLeuAspAenAspLys 161	
QY	913	TTCTCATCTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGAT---GCTAAGAGC 969	
DB	162	ValLeuIleGlnAspAenPheSerThrGlyAspGlyGlyAlaIleAenCysAlaGlySer 181	
QY	970	TTAAGGTTCAAGGAATTAGCAAGCTTTGTGCTCTTCCAGAAATATCTGCTCAAGCTGAT 1029	
DB	182	LeuLysIle----- 184	
QY	1030	GGGGAGGCTTCAAGTAGTCACCAGTTCTCTGCTATGGCTAAGGAGGCTCTTATGCCC 1089	
DB	185	-----AlaAenAsnLysSerLeuSer 191	
QY	1090	TTTGTAGCAATGTTGCAGGAGTAAGGGGGAGGATTTGCTGCTTCCAGGATGGGCGAG 1149	
DB	192	PheIleGlyAenSerSerThrArgGlyGlyAlaIleHisThr----- 206	
QY	1150	CAGGGAGTGTCATCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACT 1209	
DB	206	----- 206	
QY	1210	GGGTPAGAGTTTGTAGTGGAAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC 1269	
DB	207	-----LysAen 208	
QY	1270	GTTGCTTCTCGAATAATGGAAAAACCTGTTTCTCAACAAATGTTGCTTCTCTCTGTTTAC 1329	
DB	209	LeuThrLeuSerSerGlyGlyGluThrLeuPheGlnGlyAenThrAla----- 224	
QY	1330	ATTGCTGCTAAGCAACCAAGTGGACAGGCTTCTAATACGAGTAATAATTTACGGAGAT 1389	
DB	225	-----ProThr----- 226	
QY	1390	GGAGGAGCTATCTTCTGTGAAGAATGGTGGCAAGCAGGATCCAAATACCTCTGGATCAGTT 1449	
DB	226	----- 226	
QY	1450	TCCTTTGTAGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGA 1509	
DB	227	-----AlaAlaGlyLys----- 230	
QY	1510	GCTATTTATGCCAAAAGCTCTCGGTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAAT 1569	
DB	230	----- 230	
QY	1570	ATCGCTAATGATGGTGAGGAGGATTTATTTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCT 1629	
DB	231	-----GlyGlyAlaIleAlaIleAlaAspSerGlyThrLeuSerIleSerGly 246	
QY	1630	GATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACAGCCCAAGAGAAATCTGCC 1689	
DB	247	AspSerGlyAspIleIlePheGluGly-----AenThrIle 258	
QY	1690	GATGTTAATGGCGTAACCTGTCCTCACAGCCATTTTCATGGGATCGGAGGGAATA 1749	
DB	259	GlyAlaThrGly---ThrValSerHisSerAlaIleAspLeuGlyThrSerAlaLysIle 277	
QY	1750	ACGACATTAAGAGCTAAACGAGGCGATCAGATCTCTTTAATGATCCCATCGAGATGGCA 1809	
DB	278	ThrAlaLeuArgAlaAlaGlnGlyHisThrIleTyrPheTyrAspProIleThrVal--- 296	
QY	1810	AACGGAATAAACGAGCCAGCGAGCTCTCCAACTTCTAAATAATTAAAC----- 1857	
DB	297	ThrGlySerThrSerValAlaAspAla-----LeuAenIleAenSerProAspThr 313	
QY	1858	---GATGGTCAAGGATACACAGGGGATATTTGTTTGTCT----- 1893	
DB	314	GlyAspAenLysGluTyrThrGlyThrIleValPheSerGlyGlyLysLeuThrGluAla 333	
QY	1894	-----AATGGAAGCAGTACTTTGTACCAAAATTTCTTAAAGTGTACGATAGACAA 1935	
DB	334	GluAlaLysAspGluLysAenArgThrSerLysLeuLeuGlnAenValAlaPheLysAen 353	
QY	1936	GGAAGGATTTCTTCTCGTGAAGGCAAAATTTACTGATGAATTTCTTAAAGTGTACGACAGGT 1995	
DB	354	GlyThrValValLeuLysGlyAspValValLeuSerAlaAenGlyPheSerGlnAspAla 373	
QY	1996	GGCAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCA 2052	
DB	374	AsnSerLysLeuIleMetAspLeuGlyThrSerLeu----- 385	
QY	2053	CAACAGCCTCTGCGGCTAATCAGTTGATCAGCTTTCCTCAATCTGCAATTTGTCTCTTCT 2112	
DB	386	-----ValAlaAenThrGluSerIleGluLeuThrAenLeuGluIleAenIleAsp 402	
QY	2113	TCCTTTGTAGCAAAATGAGTTCAGATTCAGATCTCTCACTCCTCACTCCAGCGCAAGATTCT 2172	
DB	403	SerLeuArgAenGlyLysLysIle----- 410	
QY	2173	CATCTGCACTGATTGTTAGCACAACTGCTGTTCTGTCAATATTAGTGGCCTATCTTT 2232	
DB	411	-----LysLeuSerAlaAlaThrAlaGlnLysAspIleArgIleAspArgProValVal 428	
QY	2233	TTTGAGGATTTGGATGATACACAGCTTATGATAGG-----TAT 2268	
DB	429	LeuAlaIleSerAspGluSerPheTyrGlnAenGlyPheLeuAenGluAspHisSerTyr 448	
QY	2269	GATTGGCTAGTCTTAATCAAAAAATCAATGCTCTCTGAAATTTACAGTTAGGAGCTAAGCCC 2328	
DB	449	Asp-----GlyIleLeuGluLeuAspAlaGlyLysAspIle 460	
QY	2329	CCAGCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCTCAAGCTATGCTATCA 2388	
DB	461	ValIleSerAlaAspSerArgSerIleAspAlaValGlnSerPro---TyrGlyTyrGln 479	
QY	2389	GGAAGCTGGAAGCTTGGCTGGGATCCTAATAACAGCAATAATATGGTCTCTTATCTCTGAAA 2448	

```
Db 480 GlyLeuTTPThrIleAsnTTPSerThrAspAspLys-----Lys 492
Qy 2449 GCTACA-----TGACCTAAACTGGGTATATATCTGGGCTGAGCGAGTAGCTTCTTTG 2502
Db 493 AlathrValSerTTPAlaLysGlnSerPheAsnProThrAlaGluGlnGluAlaProLeu 512
Qy 2503 GTTCCAAATAGTTATGGGATCCATTTTAGATATACCATCTGCGCATTCAGCAATTCAA 2562
Db 513 ValProAsnLeuLeuTTPGlySerPheIleAspValArgSerPheGlnAsnPheIleGlu 532
Qy 2563 GCAAGTGTGGATGGGCGCTCTATTGTGCGAGGATTATGGGTTCTCGAGTTTCGAATTC 2622
Db 533 LeuGlyThrGluGlyAlaProTyrGluLysArgPheTTPValAlaGlyIleSerAsnVal 552
Qy 2623 TTCTATCATGACCGGAGTCTTTAGTCAGGATATCGGTATATATAGTGGGGTTATTCC 2682
Db 553 LeuHisArgSerGlyArgGluAsnGlnArgLysPheArgHisValSerGlyGlyAlaVal 572
Qy 2683 TTAGGAGCAACTCTACTTT---GGATCATCGATGTTTGGTCTAGCATTTACCGAAGTA 2739
Db 573 ValGlyAlaSerThrArgMetProGlyGlyAspThrLeuSerLeuGlyPheAlaGlnLeu 592
Qy 2740 TTTGTAGATCTAAAGATTATGATGTGTGCTTCCCAATCATCATGCTTGCATAGATCC 2799
Db 593 PheAlaArgAspLysAspTyrPheMetAsnThrAsnPheAlaLysThrTyrAlaGlySer 612
Qy 2800 GTTTATCTATCTACCCAAACAAGCTTTATGTGGATCCTATTTGTCGGAGATCGGTTATC 2859
Db 613 LeuArgLeu----- 615
Qy 2860 CGTGTAGTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTTCAGAGGAG 2919
Db 616 -----GlnHisAspAlaSerLeuTyrSerValVal----- 625
Qy 2920 AGCGATGTTGTTGGGATAATAACTGTCTGGTCGAGAGATTGGAGCGGATTCGGATT 2979
Db 626 -----SerIleLeuLeuGlyGlu-----GlyGlyLeuArgGlu 636
Qy 2980 GTGATTACTCCATCTAAGCTCTATTGTAAGATTGGCTCTCT-----TTCTGTGCAA 3030
Db 637 IleLeuLeuPro-----TyrValSerLysThrLeuProCysSerPheTyrGlyGln 653
Qy 3031 GCTGAGTTTCTTATGCGCATCAATCTTTTACAGAGAA----- 3072
Db 654 LeuSerTyrGlyHisThrAspHisArgMetLysThrGluSerLeuProProArg 673
Qy 3073 -----GGCATCAAGCTCGGCAATTCAGAGCGGACATCTC 3108
Db 674 ArgSerArgArgIleIleLeuLeuGlyGluAspMetSerGlyLeuGluSerTTPGluLeu 693
Qy 3109 ---CTAAATCTATCAGTTCTGTTGGAGTGAAGTTTGAATGATGTTCTAGTACACAT 3162
Db 694 GluLeuLeuLeuLysIleProAlaAlaGluAspPheSerLys---SerThrLeuHis 711

RESULT 32
G72076
polymorphic outer membrane protein h family - Chlamydoiphila pneumoniae (strain CWL029)
C/Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:99206606; PMID:10192388
A/Accession: G72076
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-978 <ARN>
A/Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000046ABB; GB:AE001629; GB:AE001363; NID
A/Experimental source: strain CWL029
C/Genetics:
A/Gene: pmp.14
C/Superfamily: Polymorphic membrane protein H family
```

```
Alignment Scores: 2.8e-40 Length: 978
Pred. No.: 702.00 Matches: 288
Score: 38.8% Conservative: 149
Best Local Similarity: 25.6% Mismatches: 419
Query Match: 8.9% Indels: 269
DB: 2 Gaps: 44

US-10-701-844-1 (1-4435) x G72076 (1-978)
Qy 405 CTTCCTTTCAATGATTCTAGCTTATTCTCTGCTCTTTAAATGGGGGGGATATGCAGC 464
Db 3 LeuSerPheLysSerSerSerPheCysLeuLeu-----AlaCysLeuCysSer 18
Qy 465 AGAAATCATGTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTA 524
Db 19 -----AlaSerCys-AlaPheAlaG1 25
Qy 525 TACTGTTATAGGAGATCGGAGTGGGACTACTGTTTTTCTGCAGAGAGTTAAACATTAA 584
Db 25 uThrArgLeuGlyGlyAsnPheValProIleThrAsnGlnGlyGluGluLeuLe 45
Qy 585 AAATCTTGAC-----AATCTATTGACAGCTTTGGCTTTTAAAGTTGTTTGGAA 632
Db 45 uThrSerAspPheValCysSerAsnPheLeuGlyAlaSerPheSerSerPheIleAs 65
Qy 633 CTTATTAGGAGTTTTACTGTTTGGAGAGGACACTCGTTGACTTTCGAGAGACATACG 692
Db 65 nSerSerSerAsnLeuSerLeuGlyLysGlyLeuSerLeuThrPheThrSerCysG1 85
Qy 693 GACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGA 752
Db 85 nAlaProThrAsnSerAsnTyrAlaLeuLeuSerAlaAlaGluThrLeu---ThrPheLy 104
Qy 753 GGGTTTTAAAGAAATTATCTCTTTTCCAAATTGCAATTCATTCTGCGTACTGCTGCTGC 812
Db 104 sAsnPheSerSerIleAsnPheThr----- 112
Qy 813 AACGACTAATAAGGTTAGCCAGACTCCGACGACAAACATCTACCCGCTAATAGTACTAT 872
Db 113 -----GlyAsnGln-----SerThrGlyLeuGlyGlyLeuI1 123
Qy 873 TTATTTCTAAACACATCTTTGTTTACTCAATAATGAGAGAGTTCTCATTCTATAGTAATTT 932
Db 123 eTyrGlyLys---AspIleValPheGlnSerIleLysAspLeuIlePheThrThrAsnAr 142
Qy 933 AGTCTCTGGAGATGGGCGAGCTATAGATGCTTAACGAGCTTAACGGTTCAAGGAATTAGCAA 992
Db 142 gVal-----AlaTyrSerProAlaSerValThrThrSerAlaThrProAl 157
Qy 993 GCTTTGTGTCTTCCAAGAAATATCTGCTCAAGCTGATGGGGGAGCTTGTCAAGTAGTCA 1052
Db 157 aile-----ThrThrValThrThrGlyAlaSerAlaLeuGlnProThrAs 172
Qy 1053 CAGTTTCTCTGCTATGCTTAACGAGGCTCTATTGCTTTTGTAGGAATGTTCAGGAGT 1112
Db 172 pSerLeuThrValGluAsnIleSerGlnSerIleLysPhePheGlyAsnLeuAlaAsnPh 192
Qy 1113 AAGAGGGGGGAGGATTGCTGCTGTTTCAGGATGGGCGAGGAGTGTCTCATCTACTTCTC 1172
Db 192 eGly-----SerAlaIleSerSe 198
Qy 1173 AACAGAAGATCCAGTAGTAGTTTTCAGAAATATCTGCG---GTAGAGTTTGTATGGGAA 1229
Db 198 rSerProThrAlaValValLysPheIleAsnAsnThrAlaThrMetSerPheSerHisAs 218
Qy 1230 CGTAGCCGAGTAGGAGGAGG---ATTACTCTACGGGACGTTGCTGCTCTCTGTAATAA 1286
Db 218 nPheThrSerSerGlyGlyGlyValIleTyrGlySerSerLeuLeuPheGluAsnAs 238
Qy 1287 TGGAAAAAACCTTGTCTTCAACAATGTTGCTTCTCTGTTTACATTTGCTGCTGAACACC 1346
```

Db 238 nSerGlyCysIleIlePheThrAlaAsnSerCysValAsnSerLeuLysGlyValThrPr 258  
QY 1347 AACAGTGCAGAGCTTCTTAATACGAGTAATAATACGAGATGGAGGAGCTATCTTCG 1406  
Db 258 oSerSerGlyThrTyAla-----LeuGlySerGlyAlaIleCysI 273  
QY 1407 TAAGAAATGTCGCAAGCAGGATCCAAATCTCTGGATCAGTTCTCTTGAT----- 1458  
Db 273 eProThrGlyThrPheGluLeuLysAsnAsnGlnGlyLysCysThrPheSerTyAsnGl 293  
QY 1459 -----GGAGAGG 1466  
Db 293 yThrProAsnAspAlaGlyAlaIleTyAlaGluThrCysAsnIleValGlyAsnGlnGl 313  
QY 1467 AGTAGTTTCTTTAGTAGCAATGCTAGCTGCTGGAAAGGGGAGCTATTATTGCCAAAA 1526  
Db 313 yAlaLeuLeuLeuLeuAspSerAsnThrAlaAlaArgAsnGlyGlyAlaIleCysAlaLysVa 333  
QY 1527 GCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATGATGGTGG 1586  
Db 333 lLeuAsnIleGlnGlyArgGlyProIleGluPheSerArgAsnArgAlaGlyGlyGl 353  
QY 1587 AGCGATTATTATTAGGA-----GAATCTGGAGAGCTCAGTTT 1622  
Db 353 yAlaIlePheIleGlyProSerValGlyAspProAlaLysGlnThrSerThrLeuThrI 373  
QY 1623 ATCTGCTGATTATGAGATATTATTTTCGATGGGAATCTTAAAGAACACAGCCAAAGAA 1682  
Db 373 eLeuAlaSerGluGlyAspIleAlaPheGlnGlyAsnMetLeuAsnThrLysPro----- 391  
QY 1683 TGTGCGGATGTTAAATGGCGTAACGTGTCTCACAAGCCATTTTCGATGGATGGGAGG 1742  
Db 392 -----GlyIle-----ArgAsnAlaIleThrValGluAlaGlyGl 403  
QY 1743 GAAATAACGACATTAAAGACTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGA 1802  
Db 403 yGluIleValSerLeuSerAlaGlnGlyGlySerArgLeuValPheTyAspProIleTh 423  
QY 1803 GATGGCAACGGAATAACGAGCGGAGTCTTCCAACTTCTTAAATTAATTAACGATGG 1862  
Db 423 rHisSerLeuProThrThrSerPro-----SerAsnLysAspIleThrIleAsn---Al 440  
QY 1863 TGAAGNATCACAGGCGATATTCTTTT----- 1890  
Db 440 aAsnGlyAlaSerGlySerValValPheThrSerLysGlyLeuSerSerThrGluLeuLe 460  
QY 1891 -----GCTAATGGAAGCAGTACTTTGTATACCAAAATGTTACGATAGACGAAGGAT 1943  
Db 460 uLeuProAlaAsnThrThrThrIleLeuLeuGlyThrValLysIleAlaSerGlyGluLe 480  
QY 1944 TGTCTTCGTGAAAGCAAAATATCAGTGAATCTCTAAGTCAGACAGGT---GGGAG 2000  
Db 480 uLysIleThrAspAsnAlaValValAsnValLeuGlyPheAlaThrGlnGlySerGlyGl 500  
QY 2001 TCTGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCCAACACCAACACAGCC 2060  
Db 500 nLeuThrLeuGlySerGlyGlyThrLeuGlyLeuAlaThr-----ProThrGlyAl 517  
QY 2061 TCTGCGCGCTAATCAGTTGATCAGCTTTTCCAACTCTGCAATTTGTCTTCTCTTTGTT 2120  
Db 517 aProAlaAlaValAsp---PheThrIleGlyLysLeuAlaPheAspProPheSerPheLe 536  
QY 2121 AGCAAAACATGACGATACGAATCTCTTACCAATCTCTCCAGCGCAAGATCTCATCTCGC 2180  
Db 536 uLysArgAspPheValSer----- 542  
QY 2181 AGTCATTTGTCAGCAACTGCTGCT-----TCTGTTACAAATTAGTGGGCTATCTTTT 2234  
Db 543 -----AlaSerValAsnAlaGlyThrLysAsnValThrLeuThrGlyAlaLeuValLe 560  
QY 2235 TGAGGATTTGATGATCAGCTTATGATAGTATGATGGCTAGTTCTAATCAAAAAA 2294  
Db 560 uAspGluHisAspValThr-----AspLeuTyAsp----- 570

QY 2295 CAATGCTCCTGAAATTTACAGTTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTCAC 2354  
Db 571 -----MetValSerLeuGlnSerProValAlaIleProIleAlaValPh 585  
QY 2355 TCTAGGGAATGAGATGCCTAAG----- 2376  
Db 585 eLysGlyAlaThrValThrLysThrGlyPheProAspGlyGluIleAlaThrProSerHi 605  
QY 2377 -TATGGCTATCAAGNAAGCTCGAAGCTTGGTGGGAT----- 2412  
Db 605 sTyGlyTyGlnGlyLysTrpSerTyThrTrpSerArgProLeuLeuLeuProAlaPr 625  
QY 2413 -----CCTAATACAGCAATAATAGTCTTATCTTATCTGAAAGCTACATGG-- 2457  
Db 625 oAspGlyGlyPheProGlyGlyProSerProSerAlaAsnThrLeuTyAlaValTrpAs 645  
QY 2458 -----ACTAAACCTGGGTATATCTCTGGGCTGAGCAGTAGCTTCTTGTGT 2504  
Db 645 nSerAspThrLeuValArgSerThrTyIleLeuAspProGluArgTyArgGlyIleVa 665  
QY 2505 TCCAATAGTTTATGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATCAA-- 2562  
Db 665 lSerAsnSerLeuTrpIleSerPheLeuGlyAsnGlnAlaPheSerAspIleLeuGlnAs 685  
QY 2563 ---GCAAGTGTGGTGGCGCTCTTATTGTGCGAGATTATGGTTCCTGGAGTTTCGAA 2618  
Db 685 pValLeuLeuIleAspHis-----ProGlyLeuSerIleThrAla----- 698  
QY 2619 TTTCTTCTATCANGACCGGATGCTTTAGT-----CAGGG 2654  
Db 699 -----LysAlaLeuGlyAlaTyValGluHisThrProArgGlnGl 712  
QY 2655 ATATCGGTATATAGTGGGGTTAT-----TCCTTAGGAGCAACCTCTTA 2699  
Db 712 yHisGluGlyPheSerGlyArgTyGlyGlyTyGlnAlaAlaLeuSerMetAsn---Ty 731  
QY 2700 CTTTGGATCATCGATTTTGGTCTAGCATTTTACCAAGATTATTTGGTAGATCTTAAAGATTA 2759  
Db 731 rThrAspHisThrThrLeuGlyLeuSerPheGlyGlnLeuTyGlyLysThrAsn--- 749  
QY 2760 TGTAGTGTCTGTTCAATCATCATCTGTTGATAGATCCGTTTATCTATCTACCAACA 2819  
Db 750 -----AlaAsnProTyAspSerArgCysSerGluGlnMetTyLeuLeuSer----- 765  
QY 2820 AGCTTTATGTCATCTTATTTCTCGA-----GATGC 2852  
Db 766 -----PhePheGlyGlnPhePheProIleValThrGlnLysSerGluAl 779  
QY 2853 GTTTATC-----CGTGTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATAT-- 2904  
Db 779 aLeuIleSerTrpLysAlaAlaTyArgTyGlyTySerLysAsnHisLeuAsnThrThrTyLe 799  
QY 2905 -----ACATTTGCAGAGGAGCGATGTTGTTGGGATAATAACTGTCTGCTGGAGA 2957  
Db 799 uArgProAspLysAlaProLysSerGlnGlyGlnTrpHisAsnAsnSerTyTyValle 819  
QY 2958 GATTGAGCGGGATTAACCGATT-----GTGATTATCTCCATCTAAGCTCTATTT 3005  
Db 819 uIleSerAlaGluHisProPheLeuAsnTrpCysLeuLeuThrArgProLeuAlaGlnAl 839  
QY 3006 GAATGATGTCGCTCTTCTGCAAGCTGAGTTTCTTATGCGGATCATGAACTCTTTTAC 3065  
Db 839 aTrpAspLeuSerGlyPheIleSerAlaGluPheLeuGlyGlyTrpGlnSerLysPheTh 859  
QY 3066 AGAGGAAGCGATCAAGCTCGGGCATTTCAAGCGGACATCTCTTAAATCTATCAGTTCC 3125  
Db 859 rGluThrGlyAspLeuGlnArgSerPheSerArgGlyLysGlyTyArgAsnValSerLeuPr 879  
QY 3126 TGTGGAGTGAAGTTTGTATGATGTTCTAGT-----ACACATCC 3164  
Db 879 oIleGly-----CysSerSerGlnTrpPheThrProPheLysLysAlaPr 894





1743	QY	GA	AAAT	TAAC	GC	AC	CA	TT	AA	GA	CT	TA	AG	CT	TA	AG	CG	GC	AT	CA	GA	T	T	C	T	T	T	T	T	A	A	T	G	A	T	C	C	C	A	T	C	G	A	1800																	
403	DB	Y	G	I	L	E	V	A	L	S	E	R	S	E	R	A	L	G	I	N	G	L	y	S	e	r	A	r	G	L	e	V	a	l	P	h	e	T	y	r	A	s	p	P	r	o	1803														
1803	QY	GA	TG	GC	CA	AA	CG	GA	AA	TA	AA	CC	AG	CG	CA	GC	CG	CA	GC	AG	CT	T	T	C	C	A	A	CT	T	C	T	T	C	T	T	C	T	T	C	T	T	A	A	1862																	
423	DB	R	H	I	S	E	R	L	e	U	P	r	o	T	h	r	S	e	r	P	r	o	-----	S	e	r	A	s	n	L	y	s	A	s	p	I	l	e	T	h	r	I	l	e	A	s	n	-----	Al	440											
1863	QY	TG	AG	GA	T	A	C	A	C	A	G	G	G	A	T	A	T	T	G	T	T	T	-----	T	G	A	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	-----	1890														
440	DB	a	E	n	G	I	y	A	l	a	S	e	r	G	l	y	S	e	r	V	a	l	P	h	e	T	h	r	S	e	r	L	y	s	G	l	y	L	e	u	S	e	r	S	e	r	T	h	r	G	l	u	L	e	u	Le	460				
1891	QY	-----	G	C	T	A	T	T	G	G	A	G	A	G	A	G	T	A	C	T	T	T	G	T	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1943												
460	DB	u	L	e	U	P	r	o	A	l	a	s	n	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	T	h	r	480					
1944	QY	TG	T	C	T	T	C	G	T	G	A	A	A	G	C	A	A	A	A	T	T	A	T	T	C	A	G	T	A	A	T	T	T	C	T	C	T	A	A	G	C	A	G	A	G	T	-----	GGGAG	2000												
480	DB	u	L	y	S	e	r	I	e	T	h	r	A	s	p	A	s	n	A	l	a	V	a	l	V	a	l	A	s	n	V	a	l	L	e	u	G	l	y	P	h	e	A	l	a	T	h	r	G	l	n	G	I	y	S	e	r	G	l	y	500
2001	QY	T	C	T	G	T	A	T	T	G	G	A	G	T	A	C	A	G	G	A	T	T	T	T	G	T	A	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2060											
500	DB	n	L	e	U	Th	r	L	e	u	G	l	y	S	e	r	G	l	y	T	h	r	L	e	u	G	l	y	L	e	u	A	l	a	T	h	-----	P	r	o	T	h	r	G	l	y	A	l	517												
2061	QY	T	C	T	G	C	G	C	T	A	A	T	C	A	G	T	T	G	A	T	C	A	G	C	T	T	C	C	A	T	T	G	C	A	T	T	T	G	C	A	T	T	T	T	T	T	T	T	T	T	T	2120									
517	DB	a	P	r	o	A	l	a	v	a	L	a	s	p	-----	P	h	e	T	h	r	I	e	G	l	y	L	y	S	e	u	A	l	a	P	h	e	A	s	p	P	r	o	P	h	e	S	e	r	P	h	e	L	536							
2121	QY	A	G	C	A	A	C	A	A	T	C	A	G	A	T	C	C	T	C	T	C	T	A	C</																																					

RESULT 34

B81593

polymorphic membrane protein H family CP0298 [imported] - Chlamydoiphila pneumoniae (str. C)  
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: B81593  
C:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, J.C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: B81593  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-978 <REA>  
 A:Cross-references: UNIPROT:Q92895; UNIPARC:UPI0000131C83; GB:AB002191; GB:AB002161; NID  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: CP0298  
 C:Superfamily: Polymorphic membrane protein H family

Alignment Scores:  
 Pred. No.: 6.19e-40 Length: 978  
 Score: 697.00 Matches: 287  
 Percent Similarity: 38.8% Conservative: 150  
 Best Local Similarity: 25.5% Mismatches: 419  
 Query Match: 8.8% Indels: 269  
 Gaps: 44

US-10-701-844-1 (1-4435) x B81593 (1-978)

```

QY 405 CTTTCTTTCAATGATCTAGCTTATTCTTGCTGCTCTTTAAATGGGGGGGATGACG 464
D 3 LeuSerPheLeuSerSerPheCysLeuLeu-----AlaCysLeuCysSer 18
QY 465 AGAAATCATGGTCTCTCAAGGAATTTACGATGGGAGAGCTTAACTGTATCATTTCCCTA 524
D 19 -----AlaSerCys-AlaPheAlaG1 25
QY 525 TACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTCTGCGAGAGAGTTAACTATAA 584
D 25 uThrArgLeuGlyGlyAsnPheValProPheLeuThrAsnGlnGlyGluGluLeuLeu 45
QY 585 AAATCTTGAC-----AATCTATTGACGCTTGCCTTTAACTGTTTGGGAA 632
D 45 uThrSerAspPheValCysAsnPheLeuGlyAlaSerPheSerSerPheLeuAs 65
QY 633 CTATTATTAGGAGTTTACTGTTTATAGGAGAGACACTCGTTGACTTTCGAGAACTAC 692
D 65 nSerSerAsnLeuSerLeuGlyGlyGlyLeuSerLeuThrPheThrSerCysG1 85
QY 693 GACTCTTCAAAATGGGCGAGCTCTAAGTAATAGCCGCTGCTGATGAGCTGTTTACTATTGA 752
D 85 nAlaProThrAsnSerAsnTy:AlaLeuLeuSerAlaAlaGluThrLeu---ThrPheLy 104
QY 753 GGGTTTTAAAGAAATATCTCTTTTCCAAATTCGAATTCATTACTTGGCTACTGCTGCTGC 812
D 104 sAsnPheSerSerIleAsnPheThr----- 112
QY 813 AACGACTAATAAGGTTAGCCAGACTCCGACGACACATCTACACCGTCTAATGTTACTAT 872
D 113 -----GlyAsnGln-----SerThrGlyLeuGlyGlyLeu11 123
QY 873 TTATTCTAAACAGATCTTTGTTACTCAATATAGAGAGTTCTCATTTCTATAGTAATTT 932
D 123 eTyTGlyLyS---AspIleValPheGlnSerIleLySAspLeuLeuPheThrThrAsnAr 142
QY 933 AOTCTCTGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAAGGTTCAAGGTTCAAGGAATTAGCA 992
D 142 gVal-----AlaTy:SerProAlaSerValThrThrSerAlaThrProAl 157
QY 993 GCTTTGTGCTTCCAAAGAAATACTGCTCAAGCTGATGGGAGCTGTCAAGTAGTACAC 1052
D 157 aile-----ThrThrValThrThrGlyAlaSerAlaLeuGlnProThrAs 172
QY 1053 CAGTTTCTCTGCTATGCTTAACGAGGCTCTATTGCTTTGTAGCGAATGTTGCGAGAGT 1112
D 172 pSerLeuThrValGluAsnIleSerGlnSerIleLySAspPheGlyAsnLeuAlaAsnPh 192
QY 1113 AAGAGGGGAGGATGCTGCTGCTTCCAGGATGGGAGGAGGTGTCATCATCTACTTC 1172
D 192 eGly-----SerAlaIleSerSe 198
QY 1173 AACAGAAGATCCAGTAGTAAGTTTTCAGAAATACTGCG---GTAGAGTTTGTATGGGAA 1229
D 198 rSerProThrAlaValValyPheIleAsnAsnThrAlaThrMetSerPheSerHisAs 218

```

```

QY 1230 CGTAGCCGAGTAGGAGAGGG---ATTACTCTACGGGAACGTTGCTTCTTCTGAATAA 1286
D 218 nPheThrSerSerGlyGlyValIleTy:GlySerSerLeuLeuPheGluAsnAs 238
QY 1287 TGGAAAAACCTTGTCTCAACATGTTGCTTCTCTGTTTACATTTGCTTAAGCAACCC 1346
D 238 nSerGlyCysIleIlePheThrAlaAsnSerCysValAsnSerLeuLySValThrPr 258
QY 1347 AACAAAGTGACAGGCTTCTTAATACGAGTAATAATTACGAGATGGAGAGCTATCTCTGTG 1406
D 258 oSerSerGlyThrTy:Ala-----LeuGlySerGlyGlyAlaIleCysI1 273
QY 1407 TAAGAATGCTCGCAAGCAGGATCCAACTCTGGATCAGTTTCTCTTGTAT----- 1458
D 273 eProThrGlyThrPheGluLeuLySAsnAsnGlnGlyLySAsnThrPheSerTy:AsnG1 293
QY 1459 -----GGAGAGGG 1466
D 293 yThrProAsnAspAlaGlyAlaIleTy:AlaGluThrCysAsnIleValGlyAsnGlnG1 313
QY 1467 AGTAGTTTCTTTAGTACCAATGTAGCTGCTGGGAAAGGGGAGCTATTTATGCCAAAAA 1526
D 313 yAlaLeuLeuLeuAspSerAsnThrAlaAlaArgAsnGlyGlyAlaIleCysAlaLySva 333
QY 1527 GCTCTCGGTTCTAACTGTGSCCTGTACAAATTTTAAAGNAATATCGTAATGATGCTGG 1586
D 333 lLeuAsnIleGlnGlyArgGlyProIleGluPheSerArgAsnArgAlaGluLySgLyG1 353
QY 1587 AGCGATTATTATGGA-----GAATCTGAGAGCTCAGTTT 1622
D 353 yAlaIlePheIleGlyProSerValGlyAspProAlaLySglnThrSerThrLeuThrI1 373
QY 1623 ATCTGCTGATTATGAGATATTTTTCGATGGGAATCTTTAAAGAACACAGCAAGAGAA 1682
D 373 eLeuAlaSerGluGlyAsnIleAlaPheGlnGlyAsnMetLeuAsnThrLySPro----- 391
QY 1693 TGCTGCCGATGTTAATGGCTAACTGTGCTCTCACAGCCATTTTCGATGGATCGGAGG 1742
D 392 -----GlyIle-----ArgAsnAlaIleThrValGluAlaGlyG1 403
QY 1743 GAAATAACGACATTAAGAGCTAAAGCAGGCGATCTCTTTTAATGATCCATCGA 1802
D 403 yGluIleValSerLeuSerAlaGlnGlyGlySerArgLeuValPheTy:AspProIleTh 423
QY 1803 GATGCAACCGAAATAACACGACGAGCTCTTCCAACTCTTCAAAATTTCAAAATTAACGATGG 1862
D 423 rHisSerLeuProThrThrSerPro-----SerAsnLySAspIleThrIleAsn---Al 440
QY 1863 TGAAGGATACACAGGGATATTGTTTTT----- 1890
D 440 aAsnGlyAlaSerGlySerValValPheThrSerLySgLyLeuSerSerThrGluLeuLe 460
QY 1891 -----GCTAATGGAAGCAGTACTTCTACCAAAATGTTACGATAGAGCAAGGAGAT 1943
D 460 uLeuProAlaAsnThrThrIleLeuLeuGlyThrValLySleAlaSerGlyGluLe 480
QY 1944 TGTTCTCGTGAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAAGCAGGT---GGAG 2000
D 480 uLySleThrAspAsnAlaValValAsnValLeuGlyPheAlaThrGlnGlySerGlyG1 500
QY 2001 TCTGTATATGAGAGCTGGAGGTACATGGGATTTGTAACTCCACCAACCCACCAACAGCC 2060
D 500 nLeuThrLeuGlySerGlyGlyThrLeuGlyLeuAlaThr-----ProThrGlyAl 517
QY 2061 TCCTCCGCTAATCAGTTGATCAGCTTTTCCAACTCTGCATTTGTCTCTTTCTCTTTGTT 2120
D 517 aProAlaAlaValAsp---PheThrIleGlyLySLeuAlaPheAspProPheSerPheLe 536
QY 2121 AGCAAAACATGAGTACGAATTCCTCTTACCAATCTCCAGCGCAAGATTTCTCATCTGCG 2180
D 536 uLySArgAspPheValSer----- 542
QY 2181 AGTCATTGGTAGCACAACTGCTGGT-----TCTGTTACAAATAGTGGGCTATCTTTT 2234

```

[illegible]

859 rGluThrGlyAspLeuGlnArgSerPheSerArgGlyLeuGlyTyrAsnValSerLeuPr 879

3126 TGTGGAGTGAAGTTTGATCGATGTTCTAGT-----ACACATCC 3164

879 oileGly-----CysSerSerGlnTrpPheThrProPheLeuLysAlaPr 894

3165 TAATAAATATAGCTTTATGCGGCTTATATCTGTGATGCTTATCGACCATCTCTGGTAC 3224

894 oSerThrLeuThrIleLysLeuAlaTyrLysProAspIleTyrArgValAsnProHisAs 914

3225 TGACACAACGCTCTATCCCATCAAGAGACATCGACAACAGATGCCCTTTCATTTAGCAAG 3284

914 nileValThrValSerAsnGlnGlusThrSerIleSerGlyAlaAsnLeuAArg 934

3285 ACATGGAGTTGTGGTTAGAGAGATCTATGTATGCTTCTTAACAAGTAAATATAGAAGTATA 3344

934 gHisGlyLeuPheValGlnIleHisAspValValAspLeuThrGluAspThrGlnAlaPh 954

3345 TGGCCATGGAGATATAGATATCGATATCGAGATGCTTCTCGAGGCTATGGTTTGATGCGAGGAG 3404

954 eLeuAsnTyrThrPheAspGlyLysAsnGlyPheThrAsnHisArgValSerThrGlyLe 974

3405 TAGAGTCGGGTTCC 3417

974 uLysSerThrPhe 978

RESULT 35

F81722

polymorphic membrane protein E/P family TC0261 [imported] - Chlamydia muridarum (strain

C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004

C:Accession: F81722

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonsay, J.; McClarty, G.; Salzberg

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: F81722

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-976 <TEXT>

A:Cross-references: UNIPROT:Q9PL47; UNIPARC:UPI0000057865; GB:AE002293; GB:AE002160; NI

A:Experimental source: strain Nigg (MoPn)

C:Genetics:

A:Gene: TC0261

Alignment Scores:	9,93e-39	Length:	976
Pred. No.:	679.50	Matches:	265
Score:	40.9%	Conservative:	150
Percent Similarity:	26.1%	Mismatches:	413
Best Local Similarity:	8.6%	Indels:	187
Query Match:	2	Gaps:	38
DB:			

US-10-701-844-1 (1-4435) x F81722 (1-976)

QY 613 CCTTTAAGTTGTTTGGGAAGTTATTAGGGAGTTTACTGTTTATGGAGAGGACATCG 672

DB 33 ProTyrHisMetThrGlyLeuPhePheProLysValAsnLeuGlyAspThrHisAsn 52

QY 673 TTGACT-----TTCGAGACATA-----CGGACTTCT 699

DB 53 LeuThrAspTyrHisLeuAspAsnLeuLysCysIleLeuAlaCysLeuGlnArgThrPro 72

QY 700 ACAATATGGGGAGCT-----CTAAGTAATAGGCGCTGCTGAT 735

DB 73 TyrGluGlyAlaAlaPheThrValThrAspTyrLeuGlyPheSerAspThrGlnLysAsp 92

QY 736 GGACTGTTTACTATTGAGGGTTTTAAAGAATTTATCTTTTCCAAATTGCAATTCATTACT 795

DB 93 GlyIlePheCys-----PheLysAsnLeuThr----- 101

QY 796 GCCGTACTGCTGTCGAACAGCTAATTAAGGTAGCCAGACTCCGACGACACATCTACA 855









Alignment Scores:	2.74e-36	Length:	938
Pred. No.:	644.00	Matches:	266
Score:	41.3%	Conservative:	160
Percent Similarity:	25.8%	Mismatches:	408
Best Local Similarity:	8.2%	Indels:	198
Query Match:	2	Gaps:	44
DB:			
US-10-701-844-1 (1-4435) x F86548 (1-938)			
QY 622	TGTTTTGGGAACCTATTAGGG---	AGTTTACTGTTT	TAGGGAGAGCACTCGTTGACT 678
DB 5	CysPheGlyMetLeuLeuProPheThrPheValLeuAlaAsnGluGlyLeuGlnLeuPro 24		
QY 679	TTGAGAACATACGGACTTCTACAAATGGGGCAGCTCTAAGT-----		-----AATAGCGCT 729
DB 25	LeuGluThrTyrlleThr-----		-----LeuSerProGluTyrlleGlnAlaAla 38
QY 730	GCTGATGGACTGTTTACTATTGAGGGTTTAAAGAAATTATCCTTT---		-----TCCAATTGCAAT 786
DB 39	ProGlnValGlyPheThrHisAsnGlnAsnGlnAspLeuAlaIleValGlyAsnHisAsn 58		
QY 787	TCATTACTTGGCGTACTCCCTGCTGCAACAGCAATAATAAGGTGACCCAGACT-----		-----837
DB 59	AspPheIleLeuAspTyrlleTyrlleTyrlleArgSerAsnGlyGlyAlaLeuThrCysLysAsn 78		
QY 838	-----	-----	-----CCGACGACAATCTACACGCTCT 861
DB 79	LeuLeuIleSerGluAsnIleGlyAsnValPhePheGluLysAsnValCysProAsnSer 98		
QY 862	AATGGTACTATTATCTTAAACAGACTCTTTTGTACTCAATAATAGAGAGTTCTCATTC 921		
DB 99	GlyGlyAlaIleTyrlleAlaGlnAsnCysThrIleSerLysAsnGlnAsnTyrlleAlaPhe 118		
QY 922	TATAGTAATTTAGTCTCTGGAGAT-----	GGGGGAGCT 954	
DB 119	ThrThrAsnLeuValSerAspAsnProThrAlaThrAlaGlySerLeuLeuGlyGlyAla 138		
QY 955	ATAGATGCTTAAGACTTAACCGTTCAAGGAATAGCAAGCTTTGTGCTTCTCCAGAAAT 1014		
DB 139	LeuPheAlaIleAsnCysSerIleThrAsnAsnLeuGlyGlnGlyThrPheValAspAsn 158		
QY 1015	ACTGCTCAAGCTGATGGGGAGCTTGCAAGTAGTCACAGTTTCTCTGTATGGCTAAC 1074		
DB 159	LeuAlaLeuAsnLysGlyGlyAlaLeuTyrlleThrGluThrAsnLeuSerIleLysAspAsn 178		
QY 1075	GAGGCTCCTATTGCTTTGTAGCGAATGTTGCA-----	GGAGTAAGAGGGGGAGG 1125	
DB 179	LysGlyProIleIleIleLysGlnAsnArgAlaLeuAsnSerAspSerLeuGlyGlyGly 198		
QY 1126	ATTGCTGCTGTTCAGTAGGGCAGCAGGAGTGTCATCTACTTCAACAGAGATCCA 1185		
DB 199	IleTyrlleSer-----	-----GlyAsnSerLeuAsnIleGlu-----	208
QY 1186	GTAGTAAGTTTTTCCAGAAATACTCGGTAGAGTTTGATGGGAACGTAGCCGAGTAGGA 1245		
DB 209	-----GlyAsnSerGlyAlaIleGlnIleThrSerAsnSerSerGlySerGly 224		
QY 1246	GGAGGGATTACTCTCTCGGGAACGTTGCTTTCTCTGTAATATGAAACCTGTTTCTC 1305		
DB 225	GlyGlyIlePheSerThrGlnThrLeuThrIleSerSerAsnLysLysLeuIleGluIle 244		
QY 1306	AACAAATGTTGCTTCTCTGTTTACATCTGCTGCTAAGCAACCAACAGTGGACAGCTTCT 1365		
DB 245	SerGlu-----	-----AsnSerAlaPheAlaAsn 252	
QY 1366	AATACGACTAATAATTACGAGATGGAGGA-----	GCTATCTCTCTGTAAGAAAT 1413	
DB 253	AsnTyrlleGlySerAsnPheAsnProGlyGlyGlyLeuThrThrThrPheCys-----	270	
QY 1414	GGTGCGCAAGCAGGATCAATAACTCTGTCAGTTTCTTTGTTGAGGAGGAGTAGTAT 1473		





```

Db      815 LeuAlaAlaPheGlnGluSerGlyAspHisAlaArgGluPhe---SerLeuHisArgPro 833
Qy      3106 CTCCTAAATCATCATGTTCTGTTGGAGTGAAGTTTGTATCGATGTTCTAGTACACATCCT 3165
Db      834 LeuThrAspValSerLeuProValGlyLe---ArgAlaSerTrpLysAsnHis 850
Qy      3166 AATAAATATAGCTTTATG---GGGCTTATATCTGTGATGCTTATTCGCACC 3213
Db      851 HisArgValProLeuValTrpLeuThrGluLeuSerTyrArgSerThrLeuTyrArgGln 870
Qy      3214 ATCTCTGCTACTGAGACACAGCTCTATCCCATCAAGACATGAGACACACAGATGCTCTT 3273
Db      871 AspProGluLeuHisLeuSerLysLeuLeuLeuSerGlnGlyThrTrpThrGlnAlaThr 890
Qy      3274 CATTTAGCAACAT---GGAGTTGTGTGTAGGATCTATG---TATGCTTCT 3321
Db      891 ProValThrTyrAsnAlaLeuGlyLeuLysValLysAsnThrMetGlnValPheProLys 910
Qy      3322 CTAAACAAGTATATAGAAGTATATGCGCATGGAAGATATGATGATGCGAGATGCTTCTCGA 3381
Db      911 ValThrLeuSerLeuAspTyr-SerAla-----AspIleSerSerThrLeu 926
Qy      3382 GCCTATGCTTTGAGTGCAGGAGTATGACTCCGGTTC 3417
Db      927 SerHisTyrLeuAsnValAlaSerArgMetArgPhe 938

RESULT 39
A:1731
polymorphic membrane protein D family TC0197 [imported] - Chlamydia muridarum (strain N1)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: AB1731
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: AB1731
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1520 <TET>
A:Cross-references: UNIPROT:Q9PLB0; UNIPARC:UPI0000057825; GB:AE002286; GB:AE002160; NID
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0197

Alignment Scores:
Pred. No.: 5,728-34 Length: 1520
Score: 611.00 Matches: 291
Percent Similarity: 37.9% Conservative: 164
Best Local Similarity: 24.2% Mismatches: 454
Query Match: 7.8% Indels: 291
DB: 2 Gaps: 55

US-10-701-844-1 (1-4435) x AB1731 (1-1520)
Qy      405 CTTTCTTTCAATGATTCT---AGCTTATTTCTGTGCTCTTTAAATGGGGGGGATATGC 461
Db      417 IleSerPheGluGlySerLysSerSerPheGlyGlyAlaAlaAlaCysGlyAsnPheSer 436
Qy      462 AGCAAAATCATGTTTCC-----TCAGAAATTTACGATGGGGA 500
Db      437 SerGluAsnAsnSerSerAlaLeuGlySerIleAspIleSerAsnAsnLeu-----Gly 454
Qy      501 GAGGTTAACTGTATCTATCTTCCCTATCTTATAGGATCCGAGTGGGACTACTGTTT 560
Db      455 AspIleSerPheLeuArgThrLeu-CysThrThrSerAspLeuGlyGlnThrAspTyrGl 474
Qy      561 TTCTGCAGAGAGATTAAACATTAAATAATCTTGACAAATCTTATGCAAGCTTTGCGCTTTAAG 620
Db      474 nGlyGlyAlaLeuPheAlaGluAsnIleSerLeuSerGluAsnAlaGlyAlaIleTh 494

```

```

Qy      621 TTGTTTGGGACCTTATTAGGAGTTTACT-----GTTTATGGAGGAGACA 668
Db      494 rPheLysAspAsnIleValLysThrPheAlaSerAsnGlyLysMetLeuGlyGlyAl 514
Qy      669 CTCCTGACTTTCGAGAACATACGAGCTCTTACAAATGGGCGAGCTCTAAGT----- 720
Db      514 alleLeuAlaSerGlyAsnValLeuIleSerLysAsnSerGlyGluLeuSerPheValGl 534
Qy      721 -AATAGCGCTGCTGATCGACTGTTTACTATATGAGGGTTTAAAGAATATCTTTTCAA 779
Db      534 YAsnAlaArgAlaProGlnAlaIleProThrArgSerSerAspGluLeuSerPhe----- 552
Qy      780 TTGCAATTCATTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGTAGCCAG---AC 836
Db      553 -----GlyAlaGlnLeuTh 557
Qy      837 TCCGACGACACATCTACACCGCTAATGTGCTACTATTATTCTAAACACAGACTTTTGT 896
Db      557 rGlnThrThrSerGlyCysSerGlyGlyAlaLeuPheGlyLys---GluValAlaIle 576
Qy      897 ACTCAATATAGGAAGTTCTCATTTATATAGTAATTTAGTCTCT----- 939
Db      576 eValGlnAsnAlaThrValPheGluGlnAsnArgLeuGlnCysGlyGlnGluTh 596
Qy      940 ----GGAGATGGGGAGCTATA---GATGCTAGAGCTTAACGTTCAAGGAATAGCAA 992
Db      596 rHisGlyGlyGlyAlaValTyrGlyMetGluSerAlaSerIleIleGly---AsnSe 615
Qy      993 GCTTTGTGCTTCCCAAGAAATATCTGCT-----CAAGCTCATGGGAGCTTTGTCA 1043
Db      615 rPheValArgPheGlyAsnAsnThrAlaValGlyAsnGlnIleSerGlyAla----- 633
Qy      1044 AGTAGTCACCAAGTTTCTGCTATGCTATGCTAAGGAGGCTCTTATTCCTTTGAGGAAATGT 1103
Db      634 -LeuLeuSerLysLysValArgLeuAlaGluAsnThrArgValAspPheSerArgAsnIle 653
Qy      1104 TGCAGGAGTAGAGGGGGAGGATGCTGCTGCTTACAGATGGGCGAGGAGGTGTCATC 1163
Db      653 eAlaThrPheCysGlyGlyAlaVal---GlnValSerAspGly-----Se 667
Qy      1164 ATCTACTTCAACAGACATCCAGTAGTAAGTTTTCAGAAATACTCGGTAGAGTTTGA 1223
Db      667 rCysGluLeuIleAsnAsnGlyTyrValLeuPheArgAspAsnArgGlyGlnThrPheGl 687
Qy      1224 TGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCAGCGGAAC-----GTTC 1274
Db      687 yGlyAlaIleSerCysLeuLysGlyAspValIleIleSerGlyAsnLysAspArgValGl 707
Qy      1275 TTTCTCGAATAT----- 1287
Db      707 uPheArgAspAsnIleValThrArgProTyrPheGluGluAsnGluGlyLysValGluTh 727
Qy      1287 ----- 1287
Db      727 rAlaAspIleAsnSerAspLysGlnGluAlaGluArgSerLeuLeuGluAsnIleGl 747
Qy      1288 -----GGAAACCTTGTCTTC-----AACATGTGCTTC 1319
Db      747 uGlnSerPheIleThrAlaThrAsnGlnThrPhePheLeuGluGluGlyLysLeuProse 767
Qy      1320 TCCTGTTTACATTCCTGCTAAGCAACCAACCAAGTGGACAG----- 1359
Db      767 rGluAlaPheIleSerAlaGluGluLeuSerLysArgGluCysAlaGlyAlaIle 787
Qy      1360 -----GCTTCTAATACG-----AGTAATAATTA 1382
Db      787 ePheAlaLysArgValTyrIleThrAspAsnLysGluProIleLeuPheSerHisAsnPh 807
Qy      1383 CGGAGAT-----GGAGAGCTATCTTCTGTAAGAATGGTCGCGAAGCAGATCAATAA 1436
Db      807 eSerAspValTyrGlyGlyAlaIlePhe-----ThrGlySerLeuGlnGluThrAspL 825
Qy      1437 CTCTGGATCAGTTCTTCTTGAT-----GGAGAGGAGTAGTTTCTTTTAG 1481

```

```
Db      825 sGlnaspValValThrProGluValValIleSerGlyAenAaspGlyAaspValIlePheSe 845
Qy      1482 TAGCAATGTAGCTGCTGGGAAA-----GGGGAGCTATTATGCG 1520
Db      845 rGlyAenAlaAlaLysHisAaspLysHisLeuProAaspThrGlyGlyAlaIleCysTh 865
Qy      1521 CAAAAGCTCTCGGTTGCT---AAGTGTGCGCTGTACAAATTTTAAGGAATATCGCTAA 1577
Db      865 rGlnAenLeuThrIleSerGlnAenAaspGlyAenValLeuPheLeuAenAaspPheAlaCy 885
Qy      1578 TGATGTGAGCGGATTTATAGGAGAATCTGAGAGCTCAGTTTATCTGCTGATATGCG 1637
Db      885 sSerGlyGlyAlaValArgIleGluAaspHisGlyGluValLeuLeuGluAlaPheGlyGl 905
Qy      1638 AGATATTATTTTCGATGGGAATCTTAAAGAACAGCAAGCAAGAGAAATGCTGCCGATTTAA 1697
Db      905 yAaspIleIlePheAenGlyAenSerPheArgAlaGln----- 918
Qy      1698 TGGCGTAACTGTCTCTCAAGCCATTTTCGATG---GGATCGGAGGAGAAATAACGAC 1754
Db      919 -----GlySerAaspAlaIleTyrPheAlaGlyLysAaspSerArgIleLysAl 934
Qy      1755 ATTAAGAGCTAAAGCAGGCGATCTCTTTAATGATCCCATCGAGATGCCAAACGG 1814
Db      934 aLeuAenAlaThrGluGlyHisAlaIleValPheGlnAaspAlaLeuValPheGlu----- 952
Qy      1815 AAATAACCAAGCCAGCGAGCTTTCCAAACTCTTAAATAATTAAC-----GATGGTGAAGG 1868
Db      953 -AenIleGluGluArgLysSerSerGlyLeuLeuValIleAenSerGlnGluAenGluGl 972
Qy      1869 ATACACAGGGATATGTTTCTCTAATGAGCAGTACTTCTACCAAAATGTTACGAT 1928
Db      972 yThrThrGlySerValArgPheLeuGlySerGluSerLysValProGlnTrpIleHisVa 992
Qy      1929 AGAGCAAGGA----- 1938
Db      992 lGlnGlnGlyLeuLeuLeuHisGlyAlaIleLeuLysSerTyrrGlyValLysGl 1012
Qy      1939 -----AGGATGTTCTTCGTGAAAGGCAAAATATATCAGTGAATTCCTTAAG 1985
Db      1012 naapProArgAlaLysIleValLeuSerAlaGlySerLysLeuLysIleLeuAaspSerGl 1032
Qy      1986 TCAGACAGGTGGAGCTGTATATGAGCTGGAGTACATGGATTTGTAATCTCCACA 2045
Db      1032 uGlnGluAenAenAla-----GluIleGlyAaspLeuGluAaspSerValAenSerGl 1049
Qy      2046 ACCACCA-----CAACAGCCTCTCGCGCTAATCAGTTGATCAGCT 2087
Db      1049 uLysThrProSerLeuTrpIleGlyLysAenAlaGlnAlaLysValProLeuValAaspIl 1069
Qy      2088 TTCCAATCTGCATTTGCTCTTTCTTTGTTAGCAACAATGCAGTTACGAATCTCTCC 2147
Db      1069 eHisThrIleSerIleAaspLeuAlaSer---PheSerSerLysAlaGlnGluThrPro-- 1087
Qy      2148 TACCAATCTCCAGCGCAAGATTCATCTCGAGTCAATGTTGGTAGCAACTGCTGGTTC 2207
Db      1088 -----GluGluAlaProGlnValIle-----ValProLysGlySe 1099
Qy      2208 TGTTACAATTAGTGGCGCTATCTTTTGGAGGATTTGGATGATACA-----GCTTATGA 2261
Db      1099 rCysValHiserGlyGluLeuSerLeuGluLeuValAasnThrThrGlyLysGlyTyrrGl 1119
Qy      2262 TAGGTATGATTTGGCTAGTTCTTAATCAAAAATCAATGTCCTCG 2304
Db      1119 uAenHisAlaLeuLeuLysAenAaspThrGlnValSerLeuMetSerPheLysGluGluAs 1139
Qy      2305 -----AAATTACGTTAGGAGTAAAGCCAGCCAGCTAATGCG 2339
Db      1139 naapGlySerLeuGluAaspLeuSerLysLeuSerVal----- 1151
Qy      2340 CCCATCAGATTTGACTCTAGGATGAGATGCTCT-----AAGTATGCGTA 2384
```

---

```
Db      1152 ---SerAaspLeuArgIleLysValSerThrProAaspIleValGluGluThrTyrrGlyHi 1170
Qy      2385 TCAAGGAAGCTGGAAGCTTGGCTGGATCTCTAATACAGCAATAATGTCTCTATATCTCT 2444
Db      1170 sMetGlyAaspTrpSerGluAla-----ThrIleGlnAaspGly-----AlaLe 1184
Qy      2445 GAAAGCTACATGAGCTAAACCTGGGTATAATCTCTGGCTGAGCAGCTAGCTCTTTGGT 2504
Db      1184 uValIleAenTrpHisProThrGlyLysLeuAaspProGlnLysAlaGlySerLeuVa 1204
Qy      2505 TCCAATAGTTTATCG-----CGATCCATTTTAGATATACGATCTCG 2546
Db      1204 lPheAenAlaLeuThrGluGluAlaValLeuSerThrLeuLysAenAlaArgIleAl 1224
Qy      2547 GCATTCAGCAATTCACGCAAGGTGGATGGCGCTCTTATCTCGAGGATTTATGGGTTTC 2606
Db      1224 aHisAenLeuThrIleGlnArgMetGlu---PheAaspTyrrSerThrAenAlaTrpGlyLe 1243
Qy      2607 TGGAGTTTTCGAATTTCTTATCATGACGCGATGCTTTA-----GGTCAGGGATATCG 2660
Db      1243 uAlaPheSerSerPheArgGluLeuSerSerGluLysLeuValSerValAaspGlyTyrrAr 1263
Qy      2661 GTATATTAGTGGGGTTATCTCTTAGGAGCAACTCTTACTTTGGATCATCG----- 2712
Db      1263 g-----GlySerTyrrIleGlyAlaSerAlaGlyIl 1273
Qy      2713 -----ATGTTGCTCTAGCATTTTACCGAAGTATTTGGTAG 2747
Db      1273 eAaspThrGlnLeuMetGluAaspPheValLeuGlyIleSerThrAlaSerPheGlyLys 1293
Qy      2748 A-----TCTAAAGATTATGTAGTGTCTTCCAATCATCATGCTTGTGATGATGATCGGT 2801
Db      1293 sMetHisSerGlnAenPhe---AaspAlaGluIleSerArgHisGlyPheValGlySerVa 1312
Qy      2802 TTATCTATCTACCCAAAGCTTTATGCGATCTTATTTGTTGCGAGATGCGGTTTATCGG 2861
Db      1312 lTyrrThr-----GlyPheLeuAlaGlyAlaTrpPhePhe 1324
Qy      2862 TGCTAGCTACGGTTTGGGAATCAGCAT-----ATGAAACCTCATATACATTTGCGAGA 2915
Db      1324 sGlyGlnTyrrSerLeuGlyGluThrHisAenAaspMetThrThrArgTyrrGlyValLeuGl 1344
Qy      2916 GGAGAGCGATGTTGCTGGGATAATAACTGTCTGCGTGGAGAGATTTGGAGCGGATTAACC 2975
Db      1344 yGluSerAenAlaThrTrpLysSerArgGlyValLeuAlaAaspAlaLeuValGluTyrrAr 1364
Qy      2976 GATGTGATTAATCTCATCT-----AAGCTCTATTGTAATGAGTTGCGTCTTCTGCGCA 3029
Db      1364 gSerLeuValGlyProAlaAargProLysPheTyrrAlaLeuHisPheAasnProTyrrValGl 1384
Qy      3030 AGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGGATCAAGCTCGGGC 3089
Db      1384 uValSerTyrrAlaSerAlaLysPheProSerPheValGluGlnGlyGlyGluAlaArgAl 1404
Qy      3090 ATTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTTGAT----- 3144
Db      1404 aPheGluGluThrSerLeuThrAasnIleThrValProPheGlyMetLysPheGluLeuSe 1424
Qy      3145 -----CGATGTTCTAGTACATCTCTAATAATATATAGCTTTATGCGGCTTATATCTG 3197
Db      1424 rPheThrLysGlyGlnPheSerGluThrAasnSerLeuGlyIleGlyCysAlaTrp----- 1442
Qy      3198 TGATGCTTATCGCACCATCTCTGCTACTGAGACACGCTCTCTCCATCCCATCAAGAGACATG 3257
Db      1443 -GluMetTyrrArgLysValGluGlyArgSerValGluLeuLeuGluAlaGlyPheAaspTr 1462
Qy      3258 GACACACAGATGCGCTTTTCATTAGCAAGACATGAGTGTGTGCTT----- 3300
Db      1462 pGluGlySerProIleAaspLeuProLysGlnGluLeuArgValAlaLeuGluAenAenTh 1482
Qy      3301 -AGAGGATCTATGATCTCTCTTAACAAGTAAATATAGATATATGCGCATGGAAGATA 3359
Db      1482 rGluTrpSerSerTyrrPheSerThrAlaLeuGlyValThrAlaPhe---CysGlyGlyPh 1501
```







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: May 13, 2006, 11:45:09 ; Search time 116.9 Seconds  
(without alignments)  
4755.534 Million cell updates/sec

Title: US-10-701-844-1  
Perfect score: 7883  
Sequence: 1 gggcaaaactctcccccg.....gcattcctaagaagaattc 4435

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues  
Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -US10701844/runat 12052006 165436 26532/app query.fasta\_1  
-Q=/abss/ABSSWEB spool/US10701844 @CGN 1 1 307 @runat 12052006 165436 26532 -NCPUP=6 -ICPU=3  
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=rapbm -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abs802p  
-USER=US10701844 @CGN 1 1 307 @runat 12052006 165436 26532 -NCPUP=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=10 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pcp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	66.8	1012	4	US-10-701-844-2
2	5267	66.8	1012	4	US-10-701-844-2
3	5267	66.8	1012	5	US-10-931-779-2
4	5131.5	65.1	1013	4	US-10-467-534-9
5	5131.5	65.1	1013	5	US-10-498-327-5
6	5131.5	65.1	1013	5	US-10-498-327-81
7	5125.5	65.0	1013	5	US-10-498-327-83
8	5123.5	65.0	1013	5	US-10-498-327-9
9	5123.5	65.0	1013	5	US-10-498-327-13
10	5123.5	65.0	1013	5	US-10-498-327-93
11	5121.5	65.0	1013	5	US-10-498-327-1

12	5121.5	65.0	1013	5	US-10-498-327-3	Sequence 3, Appli
13	5116	64.9	984	4	US-10-701-844-43	Sequence 43, Appl
14	5116	64.9	984	4	US-10-766-711-43	Sequence 43, Appl
15	5116	64.9	984	5	US-10-931-779-43	Sequence 43, Appl
16	5112.5	64.9	1013	5	US-10-498-327-17	Sequence 17, Appl
17	5106.5	64.8	1013	4	US-10-701-844-15	Sequence 15, Appl
18	5106.5	64.8	1013	4	US-10-701-844-16	Sequence 16, Appl
19	5106.5	64.8	1013	4	US-10-766-711-15	Sequence 15, Appl
20	5106.5	64.8	1013	4	US-10-766-711-16	Sequence 16, Appl
21	5106.5	64.8	1013	5	US-10-931-779-15	Sequence 15, Appl
22	5106.5	64.8	1013	5	US-10-931-779-16	Sequence 16, Appl
23	5090	64.6	1006	3	US-09-841-132-190	Sequence 190, App
24	5090	64.6	1006	5	US-10-872-155-190	Sequence 190, App
25	5084	64.5	982	3	US-08-841-132-176	Sequence 176, App
26	5084	64.5	982	5	US-10-872-155-176	Sequence 176, App
27	3336.5	42.3	670	5	US-10-197-220-169	Sequence 169, App
28	3336.5	42.3	670	6	US-11-109-468-169	Sequence 169, App
29	2552	32.4	505	4	US-10-701-844-17	Sequence 17, Appl
30	2552	32.4	505	4	US-10-766-711-17	Sequence 17, Appl
31	2552	32.4	505	5	US-10-931-779-17	Sequence 17, Appl
32	2350	29.8	458	4	US-10-701-844-36	Sequence 36, Appl
33	2350	29.8	458	4	US-10-766-711-36	Sequence 36, Appl
34	2350	29.8	458	5	US-10-931-779-36	Sequence 36, Appl
35	1735	22.0	325	4	US-10-701-844-37	Sequence 37, Appl
36	1735	22.0	325	4	US-10-766-711-37	Sequence 37, Appl
37	1735	22.0	325	5	US-10-931-779-37	Sequence 37, Appl
38	1590.5	20.2	631	3	US-09-841-132-325	Sequence 325, App
39	1590.5	20.2	631	5	US-10-872-155-325	Sequence 325, App
40	1589	20.2	1016	3	US-09-841-260-95	Sequence 95, Appl
41	1589	20.2	1016	4	US-10-007-693-95	Sequence 95, Appl
42	1589	20.2	1016	4	US-10-467-534-41	Sequence 41, Appl
43	1589	20.2	1016	4	US-10-762-058-95	Sequence 95, Appl
44	1589	20.2	1016	5	US-10-197-220-95	Sequence 95, Appl
45	1589	20.2	1016	5	US-10-498-327-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1  
US-10-701-844-2  
; Sequence 2, Application US/10701844  
; Publication No. US20040067524A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/10/701,844  
; CURRENT FILING DATE: 2003-11-04  
; PRIOR APPLICATION NUMBER: US/09/612,402B  
; PRIOR FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 1012  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-701-844-2

Alignment Scores:				
Pred. No.:	0	Length:	1012	
Score:	5267.00	Matches:	1012	
Percent Similarity:	100.0%	Conservative:	0	
Best Local Similarity:	100.0%	Mismatches:	0	
Query Match:	66.8%	Indels:	0	
DB:	4	Gaps:	0	

US-10-701-844-1 (1-4435) x US-10-701-844-2 (1-1012)

QY 382 ATGCAACGCTCTTCATAAGTTCTTCTTCAATGATTCTAGCTTATTCTTGCTCTT 441  
|||||

Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaIleValSerCysCysSer 20  
QY 442 TTAATGGGGGGATGATGAGCAGCAATCATGTTCTCTCAAGAAATTTACGATGGGAG 501  
Db 21 LeuAsnGlyGlyGlyTyraAlaAlaGluMetValProGlnGlyIleTyraAspGlyGlu 40  
QY 502 ACGTTAACTGTATCATTTCCCTATCTACTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGCAGGAGGTAAATTAATAAATCTTGACAAATCTATGAGCTTTGCGCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
QY 622 TGTTTGGGAATTTATAGGAGTTTACTGTTTATAGGAGGAGGACACTCGTGACATTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGAACATACGACTCTACAAATGGGCGAGCTCTAAGTAAATAGCGCTGTGATGGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120  
QY 742 TTTACTATTGAGGTTTAAAGAAATATCTTTTCCAAATGCAATTCATTACTTCCGCTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
QY 802 CTGCTGCTCAACGACTAATAAGGTAGCCAGACTCCGAGCAACAATCTACACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrThrSerThrProSer 160  
QY 862 AATGGTACTATTATCTTAAACAGACTTTTGTGTACTCAATAATGAGAAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
QY 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGTGTAGGGGAGCTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
QY 1042 CAAGTAGTCACAGTTCTCTGCTATGCTAACAGAGCTCTATTGCTCTTGTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240  
QY 1102 GTTCAGGAGTAAGAGGGGAGGATCTGCTGCTTCAGGATGGCGAGGAGGTGCA 1161  
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
QY 1162 TCATCTACTCAAGAGATCCAGTAGTAGTTTTCAGAAATCTGCGGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
QY 1222 GATGGAACTAGCCCGAGTAGGAGGATTTACTCTACGGGAAGCTTCTCTTCTCTG 1281  
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300  
QY 1282 AATAATGGAACCACTGTTTCTCAACAATGTGCTTCTCTGTTTACATGCTGCTAAG 1341  
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyIleAlaAlaLys 320  
QY 1342 CAACCAACAGTGGAGAGCTTCTAATACGAGTAATTAATACGAGATGGAGCTATC 1401  
Db 321 GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyGlyAspGlyGlyAlaIle 340  
QY 1402 TTCTGTAGAATGGTGGCAACAGGATCCAAATCTCTGGATCAGTTTCTTTTATGGA 1461  
Db 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly 360  
QY 1462 GAGGAGTAGTTTCTTTTAGTACCAATGTAGCTGCTGGGAAAGGGGAGCTATTATGCC 1521  
Db 361 GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyraAla 380

QY 1522 AAAAAGCTCTCGTTCTGCTAACTGTGCGCCCTGTACAAATTTTAAAGGAATATCGCTAATGAT 1581  
Db 381 LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp 400  
QY 1582 GTTGGAGCGATTTATTAGGAGATCTGGAGACTCAGTTTATCTGCTGATTTATGGAGAT 1641  
Db 401 GlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyGlyAsp 420  
QY 1642 ATTATTTTCGATCGGAATCTTAAAGAACACCCAAAGAGATGCTGCCGATGTTAATGGC 1701  
Db 421 IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly 440  
QY 1702 GTAACTGTGCTCACAAAGCCATTTGATGGATCGGAGGGGAAAATAACACATTAAGA 1761  
Db 441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg 460  
QY 1762 GCTAAAGCAGGCATCAGATTCTCTTAAATGATCCCATCGAGATGCGAAACGGAAATAAC 1821  
Db 461 AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn 480  
QY 1822 CAGCCAGCGAGTCTTCCAAACTTCTAAAAATTAAACGATGCTGAAGGATACACAGGGAT 1881  
Db 481 GlnProAlaGlnSerSerLysLeuLysIleAsnAspGlyGluGlyTyThrGlyAsp 500  
QY 1882 ATTGTTTTTCTAATGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAAAGGAGG 1941  
Db 501 IleValPheAlaAsnGlySerSerThrLeuTyGlnAsnValThrIleGluGlnGlyArg 520  
QY 1942 ATTGTTCTTCTGAAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGAGT 2001  
Db 521 IleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySer 540  
QY 2002 CTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCACAAACACCAACACAGCGCT 2061  
Db 541 LeuTyMetGluAlaGlySerThrTrpAspPheValThrProGlnProProGlnGlnPro 560  
QY 2062 CCTCCGCTAATCAGTTGATCAGCTTTTCCAAATCTGCAATTTGCTCTTCTTCTTGTGTA 2121  
Db 561 ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeu 580  
QY 2122 GCAAACTAATGAGTTCAGATCTCTACCAATCTCCAGCGCAAGATTTCTCATCTCGCA 2181  
Db 581 AlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisProAla 600  
QY 2182 GTCAATGTTAGCACCAACTGCTGTTCTGTACAATTAGTGGGCTATCTTTTTTGAGGAT 2241  
Db 601 ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAsp 620  
QY 2242 TTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTAATCAAAAATCAATGTC 2301  
Db 621 LeuAspAspThrAlaTyraAspArgTyraAspTrpLeuGlySerAsnGlnLysIleAsnVal 640  
QY 2302 CTGAAATTTACAGTTAGGAGCTAAGCCCCCAGCTAATGCCCCATCAGATTCAGTCTAGGG 2361  
Db 641 LeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAspLeuThrLeuGly 660  
QY 2362 AATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGGATCTCTAATACA 2421  
Db 661 AsnGluMetProLysTyGlyTyroGlnGlySerTrpLysLeuAlaTrpAspProAsnThr 680  
QY 2422 GCAATAATGTCCTTATCTACTCTGAAGCTACATGAGCTAAACCTGGGTATTAATCCTGG 2481  
Db 681 AlaAsnAsnGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyraAsnProGly 700  
QY 2482 CTGAGCAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCCATTTAGATATACGA 2541  
Db 701 ProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIleArg 720  
QY 2542 TCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTATTGTCGAGGATTTATGG 2601  
Db 721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeuTrp 740

```

QY 2602 GTTCTCGAGTTCGAATTTCTTCTATCATGACCGCGATGCTTTAGGTCAGGATATCGG 2661
DB 741 ValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArg 760
QY 2662 TATATTAGTGGGGTATTCTCTTAGGAGCAAACTCTCTACTTTGGATCATCGATGTTGGT 2721
DB 761 TyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPheGly 780
QY 2722 CTAGCATTTACGAAGTATTGGTAGATCTAAAGATTATGATGTGTGTCGTTCCATCAT 2781
DB 781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsnHis 800
QY 2782 CATGCTTGCATAGATCCGGTTTATCTATCTACCAACAAGCTTTATGTGATCCTATTGG 2841
DB 801 HisAlaCysIleGlySerValTyrLeuSerThrGlnGlnAlaLeuGlySerTyrLeu 820
QY 2842 TTCCGAGATCGGTTTATCCGTGCTAGCTAGCGGTTTGGGAATCAGCATATGAAAACTCA 2901
DB 821 PheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSer 840
QY 2902 TATACATTTGCAGAGAGCGATGTTGCTGGGATTAATACCTGTCTGGCTGGAGAGATT 2961
DB 841 TyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuAlaGlyGluIle 860
QY 2962 GAGCGGGATTACCGATTGTTACTTCCATCTAAGCTCTATTTGAATGAGTTGCGTCTCT 3021
DB 861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGlnLeuArgPro 880
QY 3022 TTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAAGCGCATCAA 3081
DB 881 PheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAspGln 900
QY 3082 GCTCGGGCATTCAGAGCGGACATCTCCTAAATCTATCATCTGCTGCTGTTGAGTGAAGTTT 3141
DB 901 AlaArgAlaPheLysSerGlyHisLeuLeuAsnLeuSerValProValGlyValLysPhe 920
QY 3142 GATCGATGTTCTAGTACACATCTCTAATATATAGCTTTTATGCGGCTTATATCTGTGAT 3201
DB 921 AspArgCysSerThrHisProAsnLysTyrSerPheMetAlaAlaTyrIleCysAsp 940
QY 3202 GCTTATCGCACCATCTCTGCTACTGAGACAAACGCTCTCTATCCCATCAAGAGACATGGACA 3261
DB 941 AlaTyrArgThrIleSerGlyThrGluThrLeuLeuSerHisGlnGluThrTyrThr 960
QY 3262 ACAGATGCTCTTCATTTAGCAAGACATGGAGTTGTGTTAGAGATCTATGTTGCTTCT 3321
DB 961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyrAlaSer 980
QY 3322 CTAAACAGTAATATAGAGTATATGGCCATGGAGATATGATGATGATGATGATGATGATGAT 3381
DB 981 LeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArg 1000
QY 3382 GGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
DB 1001 GlyTyrGlyLeuSerAlaGlySerArgValArgPhe 1012

```

RESULT 2

```

; US-10-766-711-2
; Sequence 2, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1012

```

```

; TYPE: PRT
; ORGANISM: Chlamydia sp.
; US-10-766-711-2

```

```

Alignment Scores:
Pred. No.: 0 Length: 1012
Score: 5267.00 Matches: 1012
Percent Similarity: 100.0% Conservative: 0
Beat Local Similarity: 100.0% Mismatches: 0
Query Match: 66.8% Indels: 0
DB: 4 Gaps: 0

```

US-10-701-844-1 (1-4435) x US-10-766-711-2 (1-1012)

```

QY 382 ATGCAAAAGCTCTTTCCATAAGTTCTTTCTTCAATGATCTAGCTTATTCTTCTGCTCT 441
DB 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTTAAATGGGGGGGATATGCAGCAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501
DB 21 LeuAsnGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGTATCATTTCCCTATATCTGTATAGGAGATCCGAGTGGGACTACTGTTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTCGCAGGAGAGTTAACTTAAATAATCTTGACAAATCTTATTCAGCTTTGCCCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTGTTTGGGAACCTTATTAGGGAGTTTTACTGTTTTAGGGAGGAGACACTCGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGACTTCTCAAAATGGGGCAGCTCTAAGTAATAGCCCTGCTGATGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATATCTTTTCCAAATTTCAATTTCACTTACCTGCGTA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTCTGTCACACACTAATAAGGCTAGCCAGACTCCGACGACCAACATCTACACCGTCT 861
DB 141 LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTCTAAAACAGACTCTTTGTTTACTCAATAATGAGAAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTCGAGNTGGGGAGCTATAGATGCTAAGAGCTTAACGCTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAAGTAGTCACCAAGTTTCTCTGCTATGGCTAAACGAGCTCTATGCTTTGTAGCGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
QY 1102 GTTTCAGGAGTAAGAGGGGAGGATTTGCTGCTTTCAGGATGGGAGGAGGAGTCTCA 1161
DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAAGATCCAGTAGTAGTAAAGTTTTTCCAGAAATACTCGGCTAGAGTTT 1221
DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTTACGGGAACGTTGCTTCCTG 1281
DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300

```





```
QY 2122 GCACAAATGCGATTACGAATCTCTACCAATCTCTCAGCGCAAGATTCATCTCATCTCGA 2181
Db 581 AlaAenAenAlaValThrAenProProThrAenProProAlaGlnAaspSerHisProAla 600
QY 2182 GTCATTGGTAGCACAACTCGTGGTCTGTGTACAATTAGTAGGGCCCTATCTTTTTCAGGAT 2241
Db 601 VallieGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAap 620
QY 2242 TTGGATGATACAGCTATGATAGGTATAGGTATGGCTAGGTCTTAATCAAAAATCAATGTC 2301
Db 621 LeuAapAapThrAlaTyAspArgTyrAspTyrLeuGlySerAasnGlnIleAenVal 640
QY 2302 CTGAATTTACAGTTAGGACCTAAGCCCGACGCTAATGCCCATCAGATTTGACTCTAGG 2361
Db 641 LeuIleGlnLeuGlnLeuGlyThrIleProAlaAenAlaProSerAaspLeuThrLeuGly 660
QY 2362 AATGAGATGCTCAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGATCCTAATACA 2421
Db 661 AenGluMetProIleTyGlyTyGlnGlySerTyrIleValLeuAlaTyrAapProAenThr 680
QY 2422 GCAATAATAGTCTCTTACTCTGAAGCTACATGAGCTAAACCTGGGTATATCTCTGGG 2481
Db 681 AlaAenAenGlyProTyrThrLeuIleValThrThrIleSerGlyTyrAasnProGly 700
QY 2482 CTTGAGCGAGTAGCTCTCTGGTTCCTCAATAGTTTATGGGATCCATTTAGATATACGA 2541
Db 701 ProGluArgValAalSerLeuValProAenSerLeuTyrGlySerIleLeuAaspIleArg 720
QY 2542 TCTGCGCATTCAGCAATCAAGCAAGTGGATGGCGCTCTTATGTCGAGGATATAGG 2601
Db 721 SerAlaHisSerAlaIleGlnAlaSerValAapGlyArgSerTyrCysArgGlyLeuTyr 740
QY 2602 GTTCTCGAGTTTCGAATTTCTTATCATGACCGCGATGCTTTTAGTCAGGATATCGG 2661
Db 741 ValSerGlyValSerAasnPheTyrHisAapArgAaspAlaLeuGlyGlnGlyTyrArg 760
QY 2662 TATATTAGTGGGGTATTCTCTAGAGCAAACTCTACTTTGGATCATCGATGTTGGT 2721
Db 761 TyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPheGly 780
QY 2722 CTAGCATTTACCGAAGTATTTGGTAGATCTAAGATTTATGTAGTGTGTCCTCAATCAT 2781
Db 781 LeuAlaPheThrGluValPheGlyArgSerLysAapTyrValValCysArgSerAenHis 800
QY 2782 CATGCTTGATAGGATCCGTTTATCTATCTACCAAGCTTTATGCGATCTTATTTG 2841
Db 801 HisAlaCysIleGlySerValTyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeu 820
QY 2842 TTTCGGAGATGCGTTTATCCGTCAGCTACGAGTTCGGGTTTGGGAATCAGCATATGAACCTCA 2901
Db 821 PheGlyAaspAlaPheIleArgAlaSerTyrGlyPheGlyAasnGlnHisMetLysThrSer 840
QY 2902 TATACATTTGCAGAGAGAGCGATGTTTCGTTGGGATAATAAATGTCCTGGCTCGAGAGAT 2961
Db 841 TyrThrPheAlaGluGluSerAapValArgTyrAapAasnAasnCysLeuAlaGlyGluIle 860
QY 2962 GGAGCGGAGATTACGATTTGATTTACTCCATCTAAGCTCTTATGAATGAGTTGCGTCT 3021
Db 861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyrLeuAasnGluLeuArgPro 880
QY 3022 TTTCGTGCAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGCGCATCA 3081
Db 881 PheValGlnAlaGluPheSerTyrAlaAapHisGluSerPheThrGluGluGlyAapGln 900
QY 3082 GCTCGGGCATCAAGAGCGGACATCTCCATAATCTATCATAGTTCTCTGTTGGAGTGAAGTTT 3141
Db 901 AlaArgAlaPheLysSerGlyHisIleLeuAasnLeuSerValProValGlyValLysPhe 920
QY 3142 GATCGATGTTCTAGTACATCTTAATAAATATAGCTTTATGGCGGCTTATATCTGTGAT 3201
Db 921 AspArgCysSerSerThrHisProAenLysTyrSerPheMetAlaAlaTyrIleCysAap 940
QY 3202 GCTTATCGCACCATCTCTGGTACTGAGACACGCTCTATCCATCATCAAGAGACATGGACA 3261
```

```
Db 941 AlaTyrArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTyrThr 960
QY 3262 ACAGATGCTTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGTATGCTTCT 3321
Db 961 ThrAapAlaPheHisLeuAlaAaHisGlyValValValArgGlySerMetTyrAlaSer 980
QY 3322 CTACACAGTATATAGAGTATATGGCCATCGAAGATATAGTATCGAGATGCTTCTCGA 3381
Db 981 LeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAaspAlaSerArg 1000
QY 3382 GGCATGCTTTTCAGTGCAGGAAGTAGATCCGGTTC 3417
Db 1001 GlyTyrGlyLeuSerAlaGlySerArgValArgPhe 1012

RESULT 4
US-10-467-534-9
; Sequence 9, Application US/10467534
; Publication No. US20040131625A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-467-534-9

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 4 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-467-534-9 (1-1013)
QY 382 ATGCAACGTCCTTTCCATAAGTCTTCTTCTTCAATGATCTAGCTTATCTTGCTCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
QY 442 TTAATATGGGGGGGATATGCAGCAAAATCATGGTTCCTCAAGGAATTTACCATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAaspGlyGlu 40
QY 502 ACGTTAACTGTATCAATTTCCCTATCTGTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAapProSerGlyThrThrValPhe 60
QY 562 TCTCGAGAGAGTTAACTTAAAAATCTTGACAAATCTTATGCAAGTCTTATGCGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAasnLeuAasnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTGTTTGGGAATTAATAGGAGGTTTTACTGTTTTAGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAasnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACATG 741
Db 101 GluAenIleArgThrSerThrAasnGlyAlaAlaLeuSerAaspSerAlaAasnSerGlyLeu 120
```



QY	742	TTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCCAAATTCGAAATTCATTACTTCCGCTA	801
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCGCTGTCGAACGACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACACCGTCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	AATGGTACTATTATTCTTAAACAGATCTTTTGTCTCAATTAATGAGAGTTCATTC	921
Db	161	AsnGlyThrIleTyfSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe	180
QY	922	TATAGTAATTTAGTCTCTCGAGATGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAATTAGCAAGCTTTGTCTTCCAAAGAAATATCTCTCAAGCTGATGGGGAGCTTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACCAAGTTTCTCTGCTATGGCTAAACGAGGCTCTTATGCTTGTAGGGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTGCAGGAGTAAGAGGGGAGGATTGCTGCTTCCAGATGGGCGAGGAGTGTC	1161
Db	241	ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGACATCCAGTAGTAAGTTTTCAGAAATACGCGGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGGAACGTAGCCGAGTAGGAGGAGGATTACTCTACGGGAACGTGCTTTCCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTyfSerTyfGlyAsnValAlaPheLeu	300
QY	1282	ATAATAGGAAAAACCTGTTTCTCAACATGTGCTTCTCTGTTTACATGCTGCTAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyfIleAlaAlaGlu	320
QY	1342	CAACCAACAGTCGACAGGCTCTTAATACGAGTAATAATTTACGGAGATGAGGAGCTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyfGlyAspGlyAlaIle	340
QY	1402	TTCTGTAAGAATCGTCGCAA---GCAGGATCCAATAACTCTGGATCAGTTTCTTGTAT	1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGAGTAGTTTCTTTTAGTACATGTTAGCTGCTGGGAAAGGGGAGCTATTTAT	1518
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyf	380
QY	1519	GCCAAAAGCTCTCGGTTGCTAACTGGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT	1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGGTCGAGCGGATTTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA	1638
Db	401	AspGlyGlyAlaIleTyfLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyfGly	420
QY	1639	GATATTATTTTCAGATGGGAATCTTTAAAGAACAGCCAAAGAGATGTCGCGATGTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGGTAACCTGTGCTCACAAGCCATTTCCATGCGATCGGAGGGAATAACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu	460
QY	1759	AGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT	1818
Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACCAGCAGCGCAGCTCTTCCAAATCTTCAAAATTTACGATGGTGAAGGATACACAGGG	1878
Db	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyfThrGly	500
QY	1879	GATATGTTTGTCTTAATGGAAGCATCTTTGTATCCAAATTTGTACATAGACGACAGA	1938
Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyfGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATGTTCTCTCGTGAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGTGGG	1998
Db	521	ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACAAACCACACAG	2058
Db	541	SerLeuTyfMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln	560
QY	2059	CCTCTCGCGCTTAATCAGTTGATCAGCTTTCCTTCCAAATCTGCATTTCTCTCTCTCTT	2118
Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580
QY	2119	TTAGCAAAACAATGCAAGTTACGAATCTCTTACCAATCTCTCCAGCGCAAGATTTCTATCCT	2178
Db	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	CGAGTCATTGGTAGCACAACTCTCTGCTCTGTTCACAAATTTAGTGGGCTATCTTTTGTAG	2238
Db	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT	2298
Db	621	AspLeuAspAspThrAlaTyfAspArgTyfAspTyfLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTCTCGAAATTAACGTAGGACTAAGCCCCAGCTTAATGCCCCATCAGATTTGACTCTA	2358
Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGATCGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTGCGGGATCCCTAAT	2418
Db	661	GlyAsnGluMetProLysTyfGlyTyfGlnGlySerTyfLysLeuAlaIleTrpAspProAsn	680
QY	2419	ACAGCAAAATAATGGTCTTATCTCTGAAAGCTACATGAGCTAAAACTGGGTATATCTCT	2478
Db	681	ThrAlaAsnAsnGlyProTyfThrLeuLysAlaThrTyfLysThrGlyTyfAsnPro	700
QY	2479	GGGCTGAGCGAGTAGCTTCTTGTTCCTCAATAGTTTATGGGATCCATTTTAGATATA	2538
Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTyfLysLeuAspIle	720
QY	2539	CGATCTCGCATTCAGCAATTCAGCAAGTGGATGGGCGCTCTTATTGTGAGGATTA	2598
Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyfCysArgGlyLeu	740
QY	2599	TGGTTTCTGGAGTTTTCGAATTTCTTCTATCATGACCGGATGCTTTAGTCTAGGATAT	2658
Db	741	TrpValSerGlyValSerAsnPheTyfHisAspArgAspAlaLeuGlyGlnGlyTyf	760
QY	2659	CGGTATATTAGTGGGGTTTCTTCTTAGGAGCAAACTCTTACTTTTGCATCATCGATGTTT	2718
Db	761	ArgTyfIleSerGlyTyfSerLeuGlyAlaAsnSerTyfPheGlySerSerMetPhe	780
QY	2719	GGTCTAGCATTTTACCAGATTTTGTGTAGATCTTAAAGATTTATGTGTGTCTTCCAAAT	2778
Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyfValValCysArgSerAsn	800
QY	2779	CATCATGCTTGGCATAGGATCCGTTTATCTATCTATCCCAACAAGCTTTATGTGATCTTAT	2838
Db	801	HisHisAlaCysIleGlySerValTyfLeuSerThrLysGlnAlaLeuCysGlySerTyf	820
QY	2839	TTGTTCCGAGATCGGTTTATCCGTAGCTAGCTAGGCTTGGGATCAGCATATGAAACC	2898
Db	821	LeuPheGlyAspAlaPheIleArgAlaSerTyfGlyPheGlyAsnGlnHisMetLysThr	840
QY	2899	TCATATACATTTTCAGAGGAGGAGCGATGTTTCTGTTGGGATAATAACTGTCTGCTGGAGAG	2958

```
Db      841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAenAsnCysLeuValGlyGlu 860
Qy      2959 ATTGAGCGGGAATTACCGAATGTGATTACTCCATCTAAGCTCTATTCAATGAGTGGGT 3018
Db      861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
Qy      3019 CCTTTCGTGCAAGCTGATTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGCAT 3078
Db      881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy      3079 CAAGCTCGGCGATTTCAAGAGCGGCATCTCCTAAATCTCATGATCCCTGTTGGAGTGAAG 3138
Db      901 GlnAlaArgAlaPheAspSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
Qy      3139 TTTCATCATGTTCTPAGTACACATCCTAATAATATAGCTTTATGGCGGCTTATATCTGT 3198
Db      921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy      3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAACGCTCCTATCCATCAAGAGACATGG 3258
Db      941 AspAlaTyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960
Qy      3259 ACACAGATGCTTTTCATTAGCAAGACATGGAGTTGTGTAGAGGATCTATGATGCT 3318
Db      961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgLysSerMetTyrAla 980
Qy      3319 TCTCTAACAAAGTAATATAGAATATATGCGCATATGCAAGATATGATATCGAGATCTTCT 3378
Db      981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy      3379 CGAGCTATGTTTGAAGTCAGGAAGTAGAGTCCGGTTC 3417
Db      1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

## RESULT 5

```
US-10-498-327-5
; Sequence 5, Application US/10498327
; Publication NO. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1013
; TYPE: PRN
; ORGANISM: Chlamydia trachomatis
US-10-498-327-5
```

## Alignment Scores:

```
Score: 0 Length: 1013
Pred. No.: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 5 Gaps: 1
```

US-10-701-844-1 (1-4435) x US-10-498-327-5 (1-1013)

```
Qy      382 ATGCAAAAGCTCTTTCCATAGTTCTTTCTTCAATGATCTAGCTTATCTTCGCTCT 441
Db      1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
Qy      442 TTAATGGGGGGGATATCAGCAGAAATCATGTTCTCTCAAGGAATTTACCATGGGAG 501
Db      21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
```

```
Qy      502 ACGTTAACTGTATCATTTCCCTATACGTATTATAGGAGATCGAGTGGAGCTACTGTTTTT 561
Db      41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy      562 TCTCAGGAGAGTTAAACATTAAAAAATCTTGACAAATCTATTGACAGCTTCCCTTTAAGT 621
Db      61 SerAlaGlyGluLeuThrLeuLysAenLeuAenSerIleAlaAlaLeuProLeuSer 80
Qy      622 TGTTTTGGGAACCTTATTAGGAGGTTTTTACTGTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
Db      81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy      682 GAGAACATACGCACTTCTCAATAGGGGAGCTCTAAGTAATAGCGCTGCTGTAGGAGCTG 741
Db      101 GluAenIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120
Qy      742 TTTTACTATGAGGGTTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTTACTTGCCTA 801
Db      121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140
Qy      802 CTGCTCTGCTCAACGCACTAATAAGGGTAGCCAGACTCCGACGACAAACATCTACACCGTCT 861
Db      141 LeuProAlaAlaThrThrAenAenGlySerGlnThrProThrThrSerThrProSer 160
Qy      862 AATGGTACTATTATTCTTAAACAGATCTTTTGTGTACTCAATATAGAGAGTTCTCATTC 921
Db      161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAenAenGluLysPheSerPhe 180
Qy      922 TATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA 981
Db      181 TyrSerAenLeuValSerGlyAspGlyAlaAlaLeuSerPheSerLeuThrValGln 200
Qy      982 GGAATTAGCAAGCTTTGTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db      201 GlyIleSerLysLeuCysValPheGlnGluAenThrAlaGlnAsePglyGlyAlaCys 220
Qy      1042 CAAGTAGTCACAGTTCTCTGCTATGCTTAACAGAGCTCTATTTGCCTTTGTAGCGAAT 1101
Db      221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240
Qy      1102 GTTCAGGAGTAAGAGGGGAGGATGCTCTGTTCAAGATGGGACAGCAGGAGTGTC 1161
Db      241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAsePglyGlnGlnGlyValSer 260
Qy      1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTTCAGAAATACTCGCGTAGAGTTT 1221
Db      261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280
Qy      1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTTACGGGAACGTTGCTTCTCTG 1281
Db      281 AspGlyAenValAlaArgValGlyGlyIleTyrSerTyrGlyAseValAlaPheLeu 300
Qy      1282 AATAATGGAAAAACCTGTTTCTCAACAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1341
Db      301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaAlaGlu 320
Qy      1342 CAACCAACAAGTGGACAGGCTTCTTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401
Db      321 GlnProThrAenGlyGlnAlaSerAenThrSerAsePheTyrGlyAsePglyGlyAlaIle 340
Qy      1402 TTCTGTAAAGATGGTGGCCAA---GCAGATCCAAATACTCTGGATAGTTTCCCTTGAAT 1458
Db      341 PheCysLysAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAsp 360
Qy      1459 GGAGAGGAGTAGTTTTTCTTTTAGTAGCAATGTAGCTGTGGGAAAAGGGGAGCTATTTAT 1518
Db      361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTyr 380
Qy      1519 GCCAAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTTTAAGGAATATCCTAAT 1578
Db      381 AlalaLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400
```

1579 GATGCTGGAGCGATTATTTAGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
1580 |||||  
401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420  
1639 GATATTATTTTCGATGGGAATCTTAAAGAAACAGCCAAAGAGAAATGCTGCCGATGTTAAT 1698  
1640 |||||  
421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
1699 GCGCTAACTGCTCTCACAAGCCATTTTCAGTGGGATCGGAGGGAATAAATACGACATTA 1758  
1700 |||||  
441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrLeu 460  
1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATATGATCCCATCGAGATGGCAACGGAAT 1818  
1819 |||||  
461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
1819 AACACGCCAGCGAGCTTCTCCAAACTTCTAAATTAACGATGGTGAAGATACACAGGG 1878  
1880 |||||  
481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyrThrGly 500  
1879 GATATTGTTTTCGCTAATCGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCA 1938  
1940 |||||  
501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
1939 AGGATTGTTCTTCGTAAGGCAAAATATCAGTCAATCTCTAAGTCAGACAGAGTGGG 1998  
1940 |||||  
521 ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACACCAACAG 2058  
2060 |||||  
541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
2059 CTTCTCGCGCTAATCAGTTGATCAGCTTTCCTCAATCGCAATTTGCTCTTCTCTCTTGG 2118  
2120 |||||  
561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
2119 TTAGCAAAACAATCAGTTAGCAATCTCTTACCAATCTCCAGCGCAAGATCTCATCCT 2178  
2180 |||||  
581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
2179 GCATCATTTGGTAGCACAACTGCTGGTTCCTGTATCAATATAGTGGGCTATCTTTTGTAG 2238  
2240 |||||  
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
2239 GATTGGATGATACAGCTTATGATAGTATGATGCTAGTGTCTAATCAAAAATCAAT 2298  
2300 |||||  
621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640  
2299 GTCTGAAATTCAGTTAGGAGCTTAAGCCCCAGCTAATGCCCATCAGATTTTGACTCTA 2358  
2360 |||||  
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
2359 GCGAATCAGATGCTTAAGTATGGCTATCAAGGAAGCTTGGAGGCTTGGTGGGATCCTTAAT 2418  
2420 |||||  
661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
2419 ACAGCAAAATATGCTCTTACTCTGAAAGCTACATGAGCTAAACTGGGTATATCCT 2478  
2480 |||||  
681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700  
2479 GGGCTGAGCGAGTATGCTTCTTGGTTCCTCAATAGTTTATGGGATCCATTTAGATATA 2538  
2540 |||||  
701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
2539 CGATCTCGCATTCAGCAATTCAGCAAGTGTGATGGGCTCTTATTTGTCAGAGATTA 2598  
2600 |||||  
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
2599 TGGGTTTCTGAGTTTCGAATTTCTTATCATGACCGCGATGCTTTAGTTCAGGATAT 2658  
2660 |||||  
741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
2659 CGGTATATTAGTGGGGTTATCTCTTAGGAGCAACTCTTCTTGTGATCATCGATGTTT 2718

761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCCAAT 2778  
2780 |||||  
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTGATCCCTAT 2838  
2840 |||||  
801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
2839 TTGTTCCGAGATCGCTTTATCCGCTAGCTACGGGTTTGGGAATCAGCATATGAACCC 2898  
2840 |||||  
821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
2899 TCATATACATTTGACAGAGAGCGATGTTCTGCGGATAATAACTGCTGCTGCGAGAG 2958  
2960 |||||  
841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
2959 ATTGAGCGGGATTAACCGATTGTGATTACTCCATCTAAGCTCTATTTGAATCAGTTGCT 3018  
2960 |||||  
861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
3019 CTTCTCGTCAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGAGCGCAT 3078  
3080 |||||  
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGlyAsp 900  
3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGAGTGAAG 3138  
3080 |||||  
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
3139 TTTGATTCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGGCGCTTATCTGT 3198  
3140 |||||  
921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
3199 GATGCTTATCGCACCATCTCTGCTACTGAGACAACTCTCTCTATCCCATCAGAGACATGG 3258  
3260 |||||  
941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
3259 ACACAGATGCTCTTTCATTTAGCACACATGAGTGTGTTAGAGCATCTATGATGCT 3318  
3260 |||||  
961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
3319 TCTTAAACAAGTAATATAGAGTATATAGCCCATGGAAGATATGAGTATCGAGATGCTTCT 3378  
3380 |||||  
981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
3379 CGAGGCTATGGTTTGTAGTGCAGAGATGAGTCCGGTTC 3417  
3380 |||||

## RESULT 6

US-10-498-327-81  
; Sequence 81, Application US/10498327  
; Publication No. US20050106162A1  
; GENERAL INFORMATION:  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis  
; FILE REFERENCE: 002441.00085  
; CURRENT APPLICATION NUMBER: US/10/498,327  
; CURRENT FILING DATE: 2004-06-10  
; PRIOR APPLICATION NUMBER: PCT/IB02/05761  
; PRIOR FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 262  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 81  
; LENGTH: 1013  
; TYPE: PRF  
; ORGANISM: Chlamydia trachomatis  
US-10-498-327-81  
Alignment Scores:

Pred. No.:	0	Length:	1013
Score:	5131.50	Matches:	983
Percent Similarity:	98.8%	Conservative:	18
Best Local Similarity:	97.0%	Mismatches:	11
Query Match:	65.1%	Indels:	1
DB:	5	Gaps:	1
US-10-701-844-1 (1-4435) x US-10-498-327-81 (1-1013)			
QY	382	ATGCAAGCTCTTCCATAAGTTCTTCTTCAATGATTTAGCTTATTTCTGCTCT	441
DB	1	MetGlnThrSerPheHisLeuPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer	20
QY	442	TAAATGGGGGGATATGCACAGAAATCATGGTTCTCAAGGATTTACAGTGGGAG	501
DB	21	LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu	40
QY	502	ACGTTAACTGATCATTTCCCTATATCTGTATAGGAGATCCGAGTGGGACTACTGTTT	561
DB	41	ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTGAGGAGGTTAAACATTTAAATACTTGACAAATCTTATTCGAGCTTTGCCCTTAA	621
DB	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
QY	622	TGTTTTGGAACTTATTAGGGAGTTTACTGTTTTAGGGAGGAGACACTCGTTGACTTTC	681
DB	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
QY	682	GAGACATACGACTTCTACAAATGGGCGAGCTCTAAGTAAATAGCGCTGATGACTG	741
DB	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
QY	742	TTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCCAAATTCGAAATTCATTTCCCGTA	801
DB	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGTGCAACGACTAATTAAGGTAGCCAGACTCCGACGACAACTACACGGTCT	861
DB	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	ATGCTACTATTATTCTAAACAGACTTTTGTGTTACTCAATAATGAGAGTTCATTC	921
DB	161	AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe	180
QY	922	TATAGTAATTAGTCTCTCGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA	981
DB	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT	1041
DB	201	GlyIleSerLysLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACAGTTTCTGCTATGCTTAACGAGCTCTTATTCCTTCTGCTAGCGAAT	1101
DB	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTGAGGAGTAGAGGGGAGGATTTGCTGTGTTTCAGGATGGGAGGAGGTGCTCA	1161
DB	241	ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTCGGTAGAGTTT	1221
DB	261	SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGAGCTAGCCGACTAGGAGGAGGATTTACTCTACGGGACGTTCTTTCCTG	1281
DB	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAAAAACCTTTGTTCTCAACAATGTTGCTTCTCTGTTTACATGCTGCTAAG	1341
DB	301	AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu	320

QY	1342	CAACCAAGTGCACAGGCTTCTTAATACGAGTAATAATACGAGATGGAGGAGCTATC	1401
DB	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle	340
QY	1402	TTCTGTAAAGATGCTGGCCAA--GCAGGATCCAAATAACTCTCGATCAGTTTCTTGTAT	1458
DB	341	PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGAGTAGTTTCTTTTAGTACAAATGCTGCTGGGAAAGGGGAGCTATTAT	1518
DB	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr	380
QY	1519	GCCAAAAGCTCTCGTTTGTAACTGTGGCCCTGTACAATTTTAAAGAAATATCGCTAAT	1578
DB	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGCTGAGCGATTTATTTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATATGGA	1638
DB	401	AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly	420
QY	1639	GATATTTTTCGATGGGAATCTTAAAGAACACCCAAAGAGAATGTCGCCGATGTTAAT	1698
DB	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGCGTAACTGTCTCCACAGCCATTTTCGATGGATCGGAGCGGAGAAATTAACGACATTA	1758
DB	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu	460
QY	1759	AGAGCTAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT	1818
DB	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACGAGCAGCGCGAGCTTCCAACTCTTAAATAATTAACGATGGTGAAGGATACACAGG	1878
DB	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyLysIleThrGly	500
QY	1879	GATATGTTTGTCTTAAGAGCAGTACTTTGTACCAAAATGTTTACGATAGACGACGA	1938
DB	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATGTTCTCTGTAAGGCAAAATTAATCAGTGAATCTCTAAGTCAGACAGGTGGG	1998
DB	521	ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACACACCACCAACAG	2058
DB	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln	560
QY	2059	CCTCTGCGGCTAACTCAGTTGATCAGCTTTCCAACTGCTGCTCTCTCTCTCTCTCTG	2118
DB	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu	580
QY	2119	TTAGCAAAACATGCAAGTTACGAATCTCTTACCAATCTCCAGCGCAAGATTCATCCT	2178
DB	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	GCAGTCATTGGTACACAACTGCTGGTCTGTGTACAAATAGTGGGCCCTATCTTTTGG	2238
DB	601	AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTTGGATGATACAGCTTATGATAGGTATGATTCGCTAGGTTCCTAATCAAAAATCAAT	2298
DB	621	AspLeuAspAspThrAlaTyrAspArgTyrAspThrLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTCTCGAAATACAGTTAGGAGCTAAGCCGCCAGCTTAATGCCCCATCAGATTGCTCTA	2358
DB	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGATGCCCTAAGTATGGCTATCAGGAAGCTGGAAGCTGCGTGGATCCTAAT	2418
DB	661	GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTrpAspProAsn	680
QY	2419	ACAGCAAAATATGGTCTTATATCTCTGAAAGCTACATGCACTAAAACTGGGTATAATCCT	2478

Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrThrLysThrGlyTyrAsnPro 700  
QY 2479 GGGCTGAGCGAGTAGCTCTTTGGTCCAAATAGCTTTATGGGATCCATTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
QY 2539 CGATCTGGCATTACGCAATTCACCAAGTGGGATGGCGCTCTTATGTCGAGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
QY 2599 TGGGTTCTGGAGTTTCGAATTTCTTATCATGACCGCATGCTTTTAGGTGAGGATAT 2658  
Db 741 TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
QY 2659 CGGTATATTAGTGGGGTTATCTTTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718  
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
QY 2719 GGTCTAGCATTTACCGAGTATTGGTAGATCTAAAGATTATGTAGTGTGCTTCCCAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
QY 2779 CATCATGCTTGATAGTAGCTGTTTATCTATCTACCCAAACAAGCTTTATGTGATCCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
QY 2839 TTGTTGCGAGATCGGTTTATCCGTGCTAGCTACGGGTTGGGAATCAGCATATGAAACC 2898  
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
QY 2899 TCATATACATTTGCAGAGAGCGATGTCGTTGGGATAAATCTGTCGCTGCGAGAG 2958  
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
QY 2959 ATTGAGCGGGATTACCGATTGGATTACTCCATCTAAGCTCTATTGTAATGAGTGGCT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGlnLeuArg 880  
QY 3019 CTTTCGTGCAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAGCGGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluArg 900  
QY 3079 CAAGCTCGGCGATTCAAGAGCGACATCTCCTAAATCTATCAGTTCTCTGTTGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGGCTTATCTGT 3198  
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
QY 3199 GATGCTTATCGCACCATCTCTGCTACTGACAGCAACGCTCTCTATCCCATCAAGAGATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTTTTCAATTTAGCAGCATGGAGTGTGCTTGTAGAGATCTATGTATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
QY 3319 TCTCTAACAGTAATATAGAGTATATGCGCATGGAAGATATCAGTATCCAGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
QY 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

## RESULT 7

US-10-498-327-83

; Sequence 83, Application US/10498327

; Publication No. US20050106162A1

; GENERAL INFORMATION:

; APPLICANT: Grandi, Guido

; APPLICANT: Ratti, Giulio  
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis  
; FILE REFERENCE: 002441, 00085  
; CURRENT APPLICATION NUMBER: US/10/498,327  
; CURRENT FILING DATE: 2004-06-10  
; PRIOR APPLICATION NUMBER: PCT/IB02/05761  
; PRIOR FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 262  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 83  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-10-498-327-83

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5125.50 Matches: 980  
Percent Similarity: 98.8% Conservative: 21  
Best Local Similarity: 96.7% Mismatches: 11  
Query Match: 65.0% Indels: 1  
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-83 (1-1013)

QY 382 ATGCAAAAGCTCTTTCCATAAGTTCTTTCTTTCAATGATTCTAGCTTATTCTTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetLeuLeuAlaTyrSerCysSer 20  
QY 442 TTAATATGGGGGGATATGCGACGAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501  
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
QY 502 AGTTAACTGATCATTTTCCCTATCTCTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGAGGAGAGCTTAACATTAATAAATCTTGCAATCTTATTCGAGCTTTGCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
QY 622 TGTGTTGGAACTTATTAGGGAGTTTTACTGTTTTAGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGACATACGACTCTTCAATATGGGAGCTCTTAAGTAATAGCGCTGCTGATGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
QY 742 TTTACTATTGAGGGTTTTAAAGAATTATCTTTTCCAATTGCAATTCATTACTTGCCGTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
QY 802 CTGCTGCTGCAACGCTTAATAAGGCTAGCCAGACTCCGACGACAACTACACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
QY 862 AATGGTACTATTATTCTTAAACAGATCTTTGTTACTCAATAATCAGAAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe 180  
QY 922 TATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGAGCTTAACCGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
QY 982 GGAATTAGCAAGCTTTGTGTTCTTCCAGAAATACTGCTCAAGCTGATGGGAGCTTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
QY 1042 CAAGTAGTCACGAGTTTCTGCTATGCTGCTTAACGAGGCTCTTATGCTTTGTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240

QY 1102 GTTCGACGAGTAAGAGGGGAGGATGCTGCTGTTCCAGGATGGCAGCAGGAGTGTC 1161  
Db ValAlaGlyValArgGlyGlyGlyLeuAlaValAlaGlnAsePGLyGlnGlnGlyValSer 260  
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTATTTTCCAGAAATACCTGCGGTAGAGTTT 1221  
Db SerSerThrSerThrGluAsePProValValSerPheSerArgAseThrAlaValGluPhe 280  
QY 1222 GATGGAAACGTAGCCCGAGTAGGAGGAGGATTTTACTCTCAGGGAACGTTGCTTTCCTG 1281  
Db AspGlyAseValAlaArgValGlyGlyGlyLeuTySerTyGlyAseValAlaPheLeu 300  
QY 1282 AATAATGAAAAACCTTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTCGCTAAG 1341  
Db AseAseGlyLeuThrLeuPheLeuAseAseValAlaSerProValTyLeAlaAlaGlu 320  
QY 1342 CAACCAACAGTCGACGCTCTAATACGAGTAATAATACGAGAGATGGAGGAGCTATC 1401  
Db GlnProThrAseGlyGlnAlaSerAseThrSerAseAseThrSerAseAseThrSerAseAse 340  
QY 1402 TTCTCTAAGAAATGGTCGCAAA---GCAGGATCCAATAACTCTGGATCAGTTTCTCTTGT 1458  
Db PheCysLeuAseGlyAlaGlnAlaAlaGlySerAseAseAseAseAseAseAseAseAse 360  
QY 1459 GGAGAGGAGTAGTTTCTTTTAGTAGCAATGATGCTGCGGAAAGGGGAGCTATTTAT 1518  
Db GlyGluGlyValAlaPhePheSerSerAseValAlaAlaGlyLeuGlyGlyAlaAlaLeuTy 380  
QY 1519 GCCAAAAGCTCTCGGTTCTAATGTCGCCCTGTACAAATTTTAAAGAAATATCGCTAAT 1578  
Db AlaLeuLeuLeuSerValAlaAseAseCysGlyProValGlnPheLeuGlyAseAseAseAse 400  
QY 1579 GATGCTGAGCGCAATTTATTTAGAGAACTCGAGAGCTCAGTTTATCTGCTGATATGGA 1638  
Db AspGlyGlyAlaLeuTyLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAseTyGly 420  
QY 1639 GATATATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAATGCTGCCGATGTTAAT 1698  
Db AspLeuLeuPheAsePGLyAseLeuLeuArgThrAlaLeuGluAseAlaAseValAse 440  
QY 1699 GCGCTAACGTGCTCTCAAGCCATTTGATGGATCGGAGGAGGAAATAACGACATTA 1758  
Db GlyValThrValSerSerGlnAlaLeuSerMetGlySerGlyGlyLeuSerThrLeu 460  
QY 1759 AGAGCTAAAGCAGGCGATCAGATTTCTTTTAATGATCCCATCGAGATGGCAACGGAAT 1818  
Db ArgAlaLeuAlaGlyHisGlnLeuPheAseAsePProLeuMetAlaAseGlyAse 480  
QY 1819 AACAGCCAGCGCAGTCTTCCAAACTTCTAATAATTAACGATGGTGAAGATACAGGG 1878  
Db AseGlnProAlaGlnSerSerGluProLeuLeuLeuAseAsePGLyGlyTyThrGly 500  
QY 1879 GATATGTTTGTGCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATACAGCAAGA 1938  
Db AspLeuValPheAlaAseGlyAseThrLeuTyGlnAseValThrLeuGlnGly 520  
QY 1939 AGGATGTTTCTTCGTAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGTGGG 1998  
Db ArgLeuValLeuArgGlyLeuAlaLeuSerValAseAseAseAseAseAseAseAseAse 540  
QY 1999 AGTCGTATATGAAGCTGGAGTGATCATGGGATTTTGTAACTTCCCAACCAACCAACAG 2058  
Db SerLeuTyMetGluAlaGlySerThrLeuAsePheValThrProGlnProGlnGln 560  
QY 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCATCTGCTTCTCTCTCTCTCTCTCTCT 2118  
Db ProProAlaAlaAseGlnLeuLeuThrLeuSerAseAseAseAseAseAseAseAseAse 580  
QY 2119 TTAGCAAAACAATGCAATGATCCTCTACCAATCCTCCAGCGCAAGATTCATCCT 2178  
Db LeuAlaAseAseAlaValThrAseProProThrAseProAlaGlnAseAseAseAseAse 600  
QY 2179 GCAGTCATTTGGTAGCAACTGCTGTTCTGTATCAATTAGTGGGCTATCTTTTGTAG 2238

Db 601 AlaLeuLeuGlySerThrAlaGlySerValThrLeuSerGlyProLeuPhePheGlu 620  
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGTCTTAATCAAAATCAAT 2298  
Db AspLeuAsePThrAlaTyAseArgTyAsePThrLeuGlySerAseGlnLeuLeuAseP 640  
QY 2299 GTCTGAAATTTACAGTTAGGAGCTAAGCCCCCAGCTAATGCCCATCAGATTTGACTCTA 2358  
Db ValLeuLeuLeuGlnLeuGlyThrGlnProSerAlaAseAlaProSerAseLeuThrLeu 660  
QY 2359 GGGATGAGATGCTTACGATAGGTATGCTATCAAGAGCTGGAAGCTTGGCTGGATCCTAAT 2418  
Db GlyAseGlnMetProLeuTyGlyTyGlnGlySerTyPheLeuAlaThrPheAseProAse 680  
QY 2419 ACAGCAATAATGCTCTTACTACTGAAAGCTACATGAGCTAATACTGGGTATAATCCT 2478  
Db ThrAlaAseAseGlyProTyThrLeuLeuAlaThrTyThrLeuTyThrGlyTyAsePro 700  
QY 2479 GGGCTGAGCAGTAGCTTCTTGGTCCAAATATGTTTATGGGATCCATTTAGATATA 2538  
Db GlyProGluArgValAlaSerLeuValProAseAseAseAseAseAseAseAseAseAse 720  
QY 2539 CGATCTGCGCATTCAGCAATTTCAAGCAAGTGTGATGGCGCTCTTATTTGTCGAGATTA 2598  
Db ArgSerAlaHisSerAlaLeuAlaSerValAsePGLyArgSerTyCysArgGlyLeu 740  
QY 2599 TGGGTTCTGAGGATTTTCGAATTTCTTATCATGACCGCGATGCTTTAGGTCAGGATAT 2658  
Db TrpValSerGlyValSerAsePheTyHisAseArgAseAlaLeuGlyGlnGlyTy 760  
QY 2659 CGGTATATTTAGTGGGGTTTCTTAGGAGCAAACTCTTACTTTTGGATCATCGATGTTT 2718  
Db ArgTyLeuSerGlyGlyTySerLeuGlyAlaAseAseAseAseAseAseAseAseAseAse 780  
QY 2719 GGTCTAGCATTTCCGAAGTATTTGGTAGACTAAGATTTAGTGTGTCGTTCCCAAT 2778  
Db GlyLeuAlaPheThrGluValPheGlyArgSerTyAseTyValValCysArgSerAse 800  
QY 2779 CATCATGCTGCATAGGATCGGTTTATCTATCTACCAACCAAGCTTTATGCGATCCTAT 2838  
Db HisHisAlaCysLeuGlySerValTyLeuSerThrTyGlnAlaLeuCysGlySerTy 820  
QY 2839 TTGTTTCGAGATGCGTTTATCGGTCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898  
Db LeuPheGlyAseAlaPheLeuArgAlaSerTyGlyPheGlyAseGlnHisMetTyThr 840  
QY 2899 TCATATACATTTGCAGAGAGCGATGCTTGGTGGGATAATACTGCTGCTGCGAGAG 2958  
Db SerTyThrPheAlaGluLeuSerAsePValArgTyPheAseAseAseAseAseAseAseAse 860  
QY 2959 ATTGAGCGGATTTACCGATTTGTGATTAATCTCATCTAAGCTCTATTTGAATGAGTTGCT 3018  
Db IleGlyValGlyLeuProLeuLeuValLeuThrProSerTyLeuTyLeuAseGlnLeuArg 880  
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATCCGATCATGAATCTTTTACAGAGAAGCGCAT 3078  
Db ProPheValGlnAlaGluPheSerTyAlaAsePHisGlySerPheThrGluGlyAse 900  
QY 3079 CAGCTCGGCAATTCAGAGCGGATCTCTTAATCTATCATGTTCTGTTGGAGTGAAG 3138  
Db GlnAlaArgAlaPheAseGlyHisLeuMetAseAseAseAseAseAseAseAseAseAse 920  
QY 3139 TTTGATCAGATGTTCTAGTACATCTCTAATAATATAGCTTTATCGCGCTTATCTGT 3198  
Db PheAsePArgCysSerSerThrHisProAseAseAseAseAseAseAseAseAseAseAse 940  
QY 3199 GATGTTTATGCAACATCTCTGGTACTGAGACAACGCTCTTATCCATCCAGAGACATGG 3258  
Db AspAlaTyArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGlnThrTy 960  
QY 3259 ACAACAGATGCTTTTCAATTTAGCAAGACATGAGTTGTTGTTAGGAGTCTATGTATGCT 3318



Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrZala 980  
Qy 3319 TCTTCAACAGTATATAGAGTATATGCGCATCGAAGATATGAGTATCGAGATCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
Qy 3379 CGAGCGTATGTTTGAAGTGCAGGAAGTAGAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerIleValArgPhe 1013

## RESULT 8

US-10-498-327-9  
; Sequence 9, Application US/10498327  
; Publication No. US20050106162A1  
; GENERAL INFORMATION:  
; APPLICANT: Grandi, Guido  
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis  
; FILE REFERENCE: 002441.00085  
; CURRENT APPLICATION NUMBER: US/10/498,327  
; CURRENT FILING DATE: 2004-06-10  
; PRIOR APPLICATION NUMBER: PCT/IB02/05761  
; PRIOR FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 262  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 9  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-10-498-327-9

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5123.50 Matches: 979  
Percent Similarity: 98.8% Conservative: 22  
Best Local Similarity: 96.6% Mismatches: 11  
Query Match: 65.0% Indels: 1  
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-9 (1-1013)

Qy 382 ATGCAAGCTCTTCCATAGTTCCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisGlyPheLeuSerMetIleLeuAlaTyrSerCysSer 20  
Qy 442 TTAATGGGGGGATATCGACAGAAATCATGGTTCCTCAAGGAATTCAGATGGGAG 501  
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
Qy 502 ACGTTAACTGATCATTTCCCTATCTGTTATAGGATCCGAGTCCGAGTACTGTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTGAGGAGAGTAAACATTAATAAATCTTGACAATCTTATGACAGCTTTGCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuIleAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80  
Qy 622 TGTGTTGGGAATTTATAGGAGTCTTCTGTTTATAGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy 682 GAGACATACGACTTCTCAAAATGGGGAGCTCTTAAGTAATAGCGTCTGCTGATGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
Qy 742 TTTTACTATTGAGGTTTAAAGAAATTTATCTTTTCCAAATTCGAATTCATTCTTCCGTA 801  
Db 121 PheThrIleGluGlyPheGlyLeuSerPheSerAsnGlyAsnSerLeuLeuAlaVal 140  
Qy 802 CTGCTGCTGCAACGACTAATAGGGTAGCCAGACTCCGACGACCAACATCTACCGCTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160

Qy 862 AATGGTACTATTTTATCTAAACAGATCTTTTGTGTACTCAATATAGAGATCTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerIleThrAspLeuLeuLeuAsnAsnGluIlePheSerPhe 180  
Qy 922 TATAGTAAATTTAGTCTCTGAGATGGGGAGCTATAGATCTAGAGCTTAACCGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaIleAspLeuThrValGln 200  
Qy 982 GGAATTAGCAAGCTTTTGTCTTCTCCAGAAAAATCTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db 201 GlyIleSerIleLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
Qy 1042 CAAGTAGTCCACGATTTCTCTGCTATCGCTAAACAGAGCTCTCTATGCTTTGTTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
Qy 1102 GTTCAGAGTAAAGGGGGAGGATGCTGCTGCTTCCAGATGGGCGAGCGAGGTGCA 1161  
Db 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTATTTTCCAGAAATACTGCGGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
Qy 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTTACGGGAAGCTTCTCTG 1281  
Db 281 AspGlyAsnValAlaAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
Qy 1282 AATAATGGAAAAACCTTGTCTTCTCAACATGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1341  
Db 301 AsnAsnGlyIleThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu 320  
Qy 1342 CAACCAACAGTGCAGAGCTTCTTAATACGAGTAGTAATAATACGAGATGGAGGAGCTATC 1401  
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
Qy 1402 TTCTGTAAAGATGGTGCACAA---GCAGGATCCCAATAACTCTGATCAGTTCTCTTGTAT 1458  
Db 341 PheCysIleAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
Qy 1459 GGAGAGGAGTAGTCTTCTTCTAGTAGCATCTAGCTGCTGGGAAAGGGGGAGCTATTTAT 1518  
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIleGlyAlaIleTyr 380  
Qy 1519 GCCAAAAGCTCTCGTTGCTAACTGTGCGCTGTACAAATTTTAAAGAAATATCCCTAAT 1578  
Db 381 AlaIleIleLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
Qy 1579 GATGTCGAGCGATTTATTTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTTATGA 1638  
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
Qy 1639 GATATTTATTTGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTCCGATGTTAAT 1698  
Db 421 AspIleIlePheAspGlyAsnLeuIleArgThrAlaIleGlyAsnAlaAlaAspValAsn 440  
Qy 1699 GCGCTAACTGTCTCTCAAGCCATTTTCGATGGGATCGGAGGAGGAAATACGACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460  
Qy 1759 AGAGCTAAAGCAGGCGCATCAGATTTCTTTTAAATGATCCATCGAGATGGCAACCGAAT 1818  
Db 461 ArgAlaIleAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
Qy 1819 AACGACCGAGCGAGCTCTTCCAAACTTCAAAAATTAACGATGGTGAAGGATACAGGG 1878  
Db 481 AsnGlnProAlaGlnSerSerGluProLeuIleAsnAspGlyGluGlyTyrThrGly 500  
Qy 1879 GATATTTGTTTGTCTAATGGAGAGCTACTTTGTACCAAAATGTTTACGATAGCAAGGA 1938  
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
Qy 1939 AGGATTTGTTCTTCGTGAAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGAGGTGG 1998



Db 521 ArgileValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
Qy 1999 AGTCTGTATATGGAAGCTGGAGTACATGGGATTTTCTAACTCCACACACACACACAG 2058  
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560  
Qy 2059 CCTCTCCGCTAATCAGTTGATCACGCTTTTCCAACTGTCATTTGCTCTCTCTCTTTG 2118  
Db 561 ProProAlaAlaGlnLeuLeuThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580  
Qy 2119 TTAGCAAAATGAGTACGATTCGATTCCTACCAATCCTCCAGCGGAGATTCATCCT 2178  
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
Qy 2179 GCAGTCAATGCTAGCACAACTGCTGTTCTGTTACAATTAGTGGCTCTATCTTTTGG 2238  
Db 601 AlaIleleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
Qy 2239 GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGGTTCTAATCAAAAATCAAT 2298  
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640  
Qy 2299 GTCTCGAAATACAGTTAGGAGCTAAGCCCGACCTAATGCCCCCATCAGATTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
Qy 2359 GGGAAATCAGATGCTTAACTAGTGGCTATCAAGAAAGCTGGAAGCTTGGTGGATCTTAAT 2418  
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
Qy 2419 ACAGCAAAATATGCTCTTATCTGAACTACATGCACTGAACTGAACTGAACTGAACTCT 2478  
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700  
Qy 2479 GGGCTGAGCAGTAGCTCTTTGGTCCAAATAGTTATGGGATCCATTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
Qy 2539 CGATCTGGCATTCAGCAATTCAGCAGATGAGTGGCGCTCTTATTGTCGAGGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
Qy 2599 TGGGTCTCTGAGTTTCGAATTTCTCTATCATCACCGCATGCTTTAGTTCAGGATAT 2658  
Db 741 TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
Qy 2659 CGGTATATTAGTGGGGTTATCTCTAGGAGCAAACTCTCTACTTTGGATCATCGATGTTT 2718  
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
Qy 2719 GGTCTAGCATTTCCGAGTATTTGGTAGATCTAAAGATTATGATGTGCTGTTCCAAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
Qy 2779 CATCATGCTGCATAGGATCGGTTTATCTATCTACCCCAAGCTTTTATGTGGATCCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
Qy 2839 TTGTTCCGAGATCGGTTTATCCGTGCTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898  
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
Qy 2899 TCATATACATTTGCAGAGAGCGATGTTCTGTTGGGATATACTGCTGCTGGCTGGAG 2958  
Db 841 SerTyrThrPheAlaGluLeuSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
Qy 2959 ATTGAGCGGATATCCGATTTGATCTTACCTCAAGCTCTATTTGAAATGAGTTGGGT 3018  
Db 861 IleGlyValGlyLeuProIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
Qy 3019 CCTTCTGGTGAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGAGCGCAT 3078

Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy 3079 CAAGCTCGGGCAATCAAGAGCGACATCTCTAAATCTATCACTTCTCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
Qy 3139 TTGTATGATGTTCTAGTACACATCCCTAATAATATAGCTTTATGGCGGCTTATATCTGT 3198  
Db 921 PheAspArgCysSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
Qy 3199 GATGCTTATCGACCATCTCTGCTACTGACACAAAGCTCTATCCCATCAAGAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACACAGATCCCTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGTATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
Qy 3319 TCTCTAAAGATATATAGAGTATATGGCCATCGGAAGATATGAGTATCGAGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
Qy 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 9  
US-10-498-327-13  
; Sequence 13, Application US/10498327  
; Publication No. US20050106162A1  
; GENERAL INFORMATION:  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis  
; FILE REFERENCE: 002441.00085  
; CURRENT APPLICATION NUMBER: US/10/498,327  
; CURRENT FILING DATE: 2004-06-10  
; PRIOR APPLICATION NUMBER: PCT/IB02/05761  
; PRIOR FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 262  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-10-498-327-13

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5123.50 Matches: 979  
Percent Similarity: 98.8% Conservative: 22  
Best Local Similarity: 96.6% Mismatches: 11  
Query Match: 65.0% Indels: 1  
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-13 (1-1013)

Qy 382 ATGCAAAAGCTCTTCCATAGTTCTTCTTCAATGATTCTAGCTTATTTCTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20  
Qy 442 TTAATGGGGGGATATGCAGCAAAATCATGTTCTCTCAAGAAATTCACATGGGAG 501  
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
Qy 502 ACGTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTCAGGAGATTAACATTAAAAAATCTTCACAAATCTTATTCGAGCTTTTGCCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80

QY	622	TGTTTTGGGAACCTTATTAGGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC	681	Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyValIleThrThrLeu	460
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100	QY	1759	AGAGCTAAAGCAGGCGCATCAGATCTCTTAAATGATCCATCGAGATGCAACAGCAAT	1818
QY	682	GAGAACATACGACTTCTACAAATGGGCGAGCTCTAAAGTAATAGCGCTCTGATGACTG	741	Db	461	ArgAlaIysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120	QY	1819	AACCAAGCAGCGCAGCTCTTCCAAACTCTTAAAAATTAACGATGGTGAAGGATACACAGG	1878
QY	742	TTTACTATTAGGGTTTTAAAGATTATCTCTTTTCCAAATGCAATTCATTCTGCGCTA	801	Db	481	AsnGlnProAlaGlnSerSerGluProLeuIysIleAsnAspGlyGluGlyThrGly	500
Db	121	PheThrIleGluGlyPheIysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140	QY	1879	GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAAGCA	1938
QY	802	CTGCGCTGCTCAACGACTAATAAGGTAGCCAGACTCCGACGACACATCTACACGCTCT	861	Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160	QY	1939	AGGATTGTTCTCTCGTGAAGGCAAAATATCATGTAATCTCTAAGTCAGACAGGTGGG	1998
QY	862	AATGGTACTATTATTCTAAAAAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTC	921	Db	521	ArgIleValLeuArgGlnIysAlaIysLeuSerValAsnSerIleuSerGlnThrGlyGly	540
Db	161	AsnGlyThrIleTyrSerIysThrAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	180	QY	1999	AGTCTGTATATCGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCAACAG	2058
QY	922	TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGACTTTACGGTTCAA	981	Db	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaIysSerLeuThrValGln	200	QY	2059	CCTCTCGCGCTAATCAGTTGATCAGCTTTCCTCAATCTGCATTTGCTCTTCTTCTTTC	2118
QY	982	GGNATTAGCAAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT	1041	Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu	580
Db	201	GlyIleSerIysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220	QY	2119	TTAGCAAAACAATGCAGTTACGAATCTCTTACCAATCTCCAGCGCAAGATTTCTCATCT	2178
QY	1042	CAAGTAGTCACCAAGTTTCTGCTATGGCTAAACGAGGCTCTTATTCCTTTGTAGCGAAT	1101	Db	581	LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240	QY	2179	CGAGTCATTGGTAGCAACAACCTGCTGTTCTGTACAAATAGTGGGCTATCTTTTGTAG	2238
QY	1102	GTTTCAGGAGTAAGAGGGGAGGATGCTGCTGTTCCAGATGGGACGAGGAGTGTC	1161	Db	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
Db	241	ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer	260	QY	2239	GATTTGGATGATACAGCTTATGATAGTATGATGCTAGTGTCTTAATCAAAAATCAAT	2298
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACTCGCGTAGAGTTT	1221	Db	621	AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnIysIleAsp	640
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280	QY	2299	GTCTCGAAATACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA	2358
QY	1222	GATGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTCTCCTG	1281	Db	641	ValLeuIysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300	QY	2359	GGGAATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTCGCTGGGATCTAAT	2418
QY	1282	AATAATGGAAAACTTTGTTCTCAAAATGTTGCTTCTCTGTTTACATGCTGCTAAG	1341	Db	661	GlyAsnGluMetProIysTyrGlyTyrGlnGlySerTyrIleValLeuAlaTyrAspProAsn	680
Db	301	AsnAsnGlyIysThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu	320	QY	2419	ACAGCAAAATAATGGTCTTATCTCTCAAGAGCTACATGGACTAAACTGGGTATATCTCT	2478
QY	1342	CAACCAACAAGTGGACAGGCTTCTAATACGATTAATAATTACGAGATGAGGAGCTATC	1401	Db	681	ThrAlaAsnAsnGlyProTyrThrLeuIysAlaThrTriphThrIysThrGlyTyrAsnPro	700
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle	340	QY	2479	GGGCTGAGCGAGTAGTCTTCTTGGTTCCAAATAGTTATGGGATCCATTTTAGATATA	2538
QY	1402	TTCTGTAAAGATGGTGGCAAA---GCAGGATCCAATAACTCTGGATCAGTTCCTTTGAT	1458	Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle	720
Db	341	PheCysIysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360	QY	2539	CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGGCGCTCTTATTGTCGAGGATTA	2598
QY	1459	GGAGAGGAGTAGTATTTCTTTAGTAGCAATGTAGCTGCTGGGAAGGGGAGCTATTTAT	1518	Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIysGlyAlaIleTyr	380	QY	2599	TGGGTTTCTCGAGTTCGAAATTTCTTCTATCATGACCGCATGCTTTAGTCTAGGAGTAT	2658
QY	1519	GCCAAAAGCTCTCGGTGCTAAGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT	1578	Db	741	TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760
Db	381	AlaIysIysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400	QY	2659	CGGTATATTAGTGGGGTTATCTCTTAGGAGCAATCTCTTACTTTGGATCATCGATCTTT	2718
QY	1579	GATCGTCGAGGATTTTATAGGAGAATCTGAGAGCTCAGTTTATCTGCTGATTATGGA	1638	Db	761	ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780
Db	401	AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly	420	QY	2719	GGTCTAGCATTTACCGAAGTATTGTTAGATCTAAAGATTATGATGTTGCTGTTCAAT	2778
QY	1639	GATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAGAGAAATGCTGCGGATTTAAT	1698	Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerIysAspTyrValValCysArgSerAsn	800
Db	421	AspIleIlePheAspGlyAsnLeuIysArgThrAlaIysGluAsnAlaAlaAspValAsn	440	QY	2779	CATCATCTTGGCATAGGATCCGTTTATCTATCTACCAACAAGCTTTTATGTTGATCTTAT	2838
QY	1699	GGCGTAATGTGTCTCTCAAGCCATTTTCGATGGGATCGGGAGGAAAAATACGACATTA	1758				

Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
Qy 2839 TTGTTCCGAGATGCGTTTATCCGCTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898  
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
Qy 2898 TCATATACATTCGACAGAGAGCGATGTCGTTGGGATATTAATCTGCTGGCGAGAG 2958  
Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
Qy 2959 ATTGAGCGGATTAACCATGTTGATTAATCTCAAGCTCTATTGAATGAGTTGGT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGlnLeuArg 880  
Qy 3019 CTTTTCGTCGAGCTGAGTTTCTTATCCCGATCATGAATCTTTTACAGAGAAAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy 3079 CAAGCTCGGCATTCAGAGCGGACATCTCCTAAATCTCATGTTCTCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
Qy 3139 TTTCATCATGTTCTAGTACACATCTTAATAATATAGCTTTATGCGCGCTTATATCTGT 3198  
Db 921 PheAspArgCysSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
Qy 3199 GATGCTTATCGCACATCTCTGGTACTGAGACACCGCTCCTATCCCATCAAGAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACAACAGATGCTTCATTTAGCAGACATGGAGTTGGTTAGAGGATCATGTATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
Qy 3319 TCTCAACAAGTAATATAGATATATGCCATGGAAGATAGATGAGATCGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
Qy 3379 CGAGCTATGTTGAGTCAGGAAGTAGAGTCCGTTCT 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

## RESULT 10

US-10-498-327-93

; Sequence 93, Application US/10498327

; Publication NO. US20050106162A1

; GENERAL INFORMATION:

; APPLICANT: Grandi, Guido

; APPLICANT: Ratti, Giulio

; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis

; FILE REFERENCE: 002441.00085

; CURRENT APPLICATION NUMBER: US/10/498,327

; CURRENT FILING DATE: 2004-06-10

; PRIOR APPLICATION NUMBER: PCT/IB02/05761

; PRIOR FILING DATE: 2002-12-12

; NUMBER OF SEQ ID NOS: 262

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 93

; LENGTH: 1013

; TYPE: PRT

; ORGANISM: Chlamydia trachomatis

US-10-498-327-93

Alignment Scores:

Pred. No.: 0 Length: 1013

Score: 5123.50 Matches: 979

Percent Similarity: 98.8% Conservative: 22

Best Local Similarity: 96.6% Mismatches: 11

Query Match: 65.0% Indels: 1

DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-93 (1-1013)

Qy 382 ATGCACACGCTTCCATAGTTCTTCTTCAATGATCTTAGCTTATTCTTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20  
Qy 442 TTAATATGGGGGATATGACGACGAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501  
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40  
Qy 502 ACGTTAACTGATCATTTCCCTATCTACTGTTATAGAGATCCGAGTGGAGCTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTCAGGAGAGTTAAACATTAATAAATCTTGACAAATCTATTCGACGTTTCCCTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80  
Qy 622 TGTTTTGGGAATCTATTAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy 682 GAGAACATACGACCTTCTACAAATGGGCGAGCTCTTAAGTAATAGCGCTGCTGATCGACTG 741  
Db 101 GluAsnIleAlaGThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
Qy 742 TTTACTATTGAGGCTTTTAAAGAATTATCTCTTTTCCAAATTCGAATTCATTACTTCCCTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
Qy 802 CTGCTCTGCTCAACGACCTAATAAGGTAGCCAGACTCCGACGACCAACATCTACACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160  
Qy 862 AATGCTACTATTATTCTTAAACAGACTCTTTTGTGTACTCAATAATAGAGAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGCTTAACGGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
Qy 982 GGAATTAGCAAGCTTTGTGCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
Qy 1042 CAAGTAGTACACAGTTCTCTGCTATGCTTAACGAGGCTCTTATTCCTTTGTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
Qy 1102 GTTCGAGGATGAGAGGGGAGGATTGCTGCTGTTTCAGGATGGCAGCGGAGTGTCA 1161  
Db 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlyValSer 260  
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTCGCGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
Qy 1222 GATGGAACTGATCCCGAGTAGGAGGAGGATTTACTCTTACGGGAACTGCTTTCTCTG 1281  
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
Qy 1282 AATATGGAAACCTTGTCTTCTCACAATGTTGCTTCTCTGTTTACATTTGCTGCTAAG 1341  
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu 320  
Qy 1342 CAACCAACAGTGGACAGGCTTCTAATACAGTAATAATATACGAGATGGAGGAGTATC 1401  
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle 340  
Qy 1402 TTCTGTAAAGATGGTGGCGAA---GCAGGATCCAATAACTCTGGATACGATTTCTTTGAT 1458  
Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
Qy 1459 GGAGAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518

Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyGlyGlyAlaIleTyr 380  
 QY 1519 GCCAAAGCTCTCGGTTCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578  
 Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
 QY 1579 GATGGTGGAGCAATTTATTTAGGAGAACTCTGGAGAGCTCAGTTTATCTCTGCTGATTATGGA 1638  
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420  
 QY 1639 GATATTATTTCGATCGGGAATCTTAAAGAAACAGCCAAAGAGAGAAATCTCCGATGTTAAT 1698  
 Db 421 AspilellePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
 QY 1699 GCGTAACTGTGTCTCAAGCCATTTCGATGGGATCGGGAGGGGAAATAAAGCAATTA 1758  
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
 QY 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTATGATCCCATCGAGATGGCAACGGAAAT 1818  
 Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
 QY 1819 AACCCAGCAGCGAGCTTCCAAACTCTTAAATAATTAAACGATGGTGAAGGATACACAGGG 1878  
 Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500  
 QY 1879 GATATTGTTTCTTAATGGAAGCAGTACTTTCTACCAAAATTTTACGATAGAGCAAGGA 1938  
 Db 501 AspileValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
 QY 1939 AGGATTGTTCTCGTGAAGGCAAAATTTATCAGTGAATCTCTAGTCAAGCAGAGTGGG 1998  
 Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
 QY 1999 AGTCTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCAAACACCAACACACAG 2058  
 Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
 QY 2059 CCTCTCGCCCTAATCAGTTGTATCAGCTTTCCAACTTCGATTTGCTCTCTCTTTG 2118  
 Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580  
 QY 2119 TTAGCAAAACATGAGTTAGCAATCTCTCCACCAATCTCCAGGCAAGATTTCTCATCCT 2178  
 Db 581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
 QY 2179 GCAGTCAATGTAGCAACCTGCTGTTCTGTACAAATTAAGTGGCCCTATCTTTTGTAG 2238  
 Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
 QY 2239 GATTGGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298  
 Db 621 AspLeuAspAspThrAlaTyrAspAspTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640  
 QY 2299 GTCTGAAATTAAGTTAGGACCTAAGCCCCAGCTAATGCCCATCAGATTGCACTTA 2358  
 Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
 QY 2359 GGGATCAGATGCTTAAGTATGGCTATCAAGGAGCTGGAAGCTTGGTGGATCCTTAAT 2418  
 Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaIlePheProAsn 680  
 QY 2419 ACAGCAAAATATGCTCTTACTCTGAAAGCTACATGGACTAAAGCTGGGTATTAATCCT 2478  
 Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrIleThrLysThrGlyTyrAsnPro 700  
 QY 2479 GGGCTCAGCAGTGTCTCTTGGTCCAAATAGTTTATGGGATCCATTTAGATATA 2538  
 Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720  
 QY 2539 CGATCTCGCAATTCAGCAATTCAGCAAGTGTGGATGGGCTCTTATTGTGAGGATTA 2598

Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
 QY 2599 TGGGTTTCTGGAGTTTCCGAATTTCTTATCATGACCGCGATGCTTTTAGGTGAGGATAT 2658  
 Db 741 trpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
 QY 2659 CGGTATATTAGTGGGGTTATTCTTTAGGAGCAAACTCTCTACTTTGGATCATCGATGTTT 2718  
 Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
 QY 2719 GGTCTAGCAATTTACCGAAGTATTTGGTAGATCTTAAGATTATGTAGTGTGCTGTTCCAAT 2778  
 Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
 QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTTATCTACCCAAACAGCTTTTATGTGATCCTAT 2838  
 Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
 QY 2839 TTGTTTCGGAGATCGGTTTATCCGCTAGCTACGCTTGGGAAATCAGCATATGAAAAACC 2898  
 Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
 QY 2899 TCATATACATTTTCGAGAGAGAGCGATGTTCTGTTGGGATAATACTGTCTGGCTGAGAG 2958  
 Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
 QY 2959 ATTGGAGCGGATTTACCGATTCTGATTACTCCATCTTAAGCTCTTATTTGAATGAGTTGCGT 3018  
 Db 861 IleGlyValGlyLeuProIleIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
 QY 3019 CTTTTCGTCGCAAGCTGAGTTTCTTATGCGCATCATGAAATCTTTTACAGAGGAAGCGAT 3078  
 Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
 QY 3079 CAAGCTCGGCGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGAGTGAAG 3138  
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
 QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAAATATAGCTTTTATGCGGCTTATATCTGT 3198  
 Db 921 PheAspArgCysSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
 QY 3199 GATGCTTATCGCACCATCTCTGCTAGACAAACGCTCTCTATCCCATCAAGAGACATGG 3258  
 Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
 QY 3259 ACAACAGATGCTTTTCATTTAGCAACATGAGGATGTTGTTGTTAGAGATCTATGATGCT 3318  
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
 QY 3319 TCTCTAACAGTAATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTTCT 3378  
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
 QY 3379 CGAGGCTATGTTGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417  
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 11  
 US-10-498-327-1  
 ; Sequence 1, Application US/10498327  
 ; Publication No. US20050106162A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Grandi, Guido  
 ; APPLICANT: Ratti, Giulio  
 ; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis  
 ; FILE REFERENCES: 002441.00085  
 ; CURRENT APPLICATION NUMBER: US/10/498.327  
 ; CURRENT FILING DATE: 2004-06-10  
 ; PRIOR APPLICATION NUMBER: PCT/IB02/05761  
 ; PRIOR FILING DATE: 2002-12-12  
 ; NUMBER OF SEQ ID NOS: 262  
 ; SOFTWARE: PatentIn version 3.2



```

Db      641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660
QY      2359 GGAATGAGATCCCTAAGTATGGCTATCAAGAGCTGAAGCTTGGCTGGATCCTAAT 2418
Db      661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAen 680
QY      2419 ACAGCAAAATAGTGGCTTATCTCTGAAGCTACATGAGCTAAAGCTGGGTATATCTCT 2478
Db      681 ThrAlaAenAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAenPro 700
QY      2479 CGGCTCAGCAGTAGTCTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTATAGATA 2538
Db      701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
QY      2539 CGATCTGCCATTCAGCAATTCAGCAAGTGGATGGCGGCTCTTATTGTCGAGGATTA 2598
Db      721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY      2599 TGGGTTCTGGAGTTTCGAATTTCTCTATCATGACCGGATGCTTTAGTTCAGGGATAT 2658
Db      741 TrpValSerGlyValSerAenPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY      2659 CGGTATATTAGTGGGGTTTATCTCTTAGGAGCAAACTCTCTACTTTGGATCATCGATGTTT 2718
Db      761 ArgTyrIleSerGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
QY      2719 GGTCTAGCATTTACCAAGATTTGGTAGATCTAAAGATTTAGTAGTGTGTTCCAT 2778
Db      781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800
QY      2779 CATCATGCTGCATAGATCCGTTTATCTCTATCTACCCCAACAGCTTTTATGTGATCCTAT 2838
Db      801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY      2839 TTGTTCCGAGATCGGTTTATCCGTCTAGCTACGCGTTTGGGAATCGAGATATGAAACC 2898
Db      821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
QY      2899 TCATATACATTTCCAGAGAGCGATGTTCTGGTGGGATTAATCACTGCTGGCTGGAGAG 2958
Db      841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAenAsnCysLeuValGlyGlu 860
QY      2959 ATTGGAGCGGATACCGATTGATCTGATCTCACTCAATCTTAAGTCTATTGATGATGTCGT 3018
Db      861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
QY      3019 CTTTTCGTCAAGCTCAGATTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGAT 3078
Db      881 PropheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY      3079 CAAGCTCGGCGATTCAGAGCGGACATCTCTAAATCTATCACTTCTGTTGGAGTGAAG 3138
Db      901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
QY      3139 TTTGATCGATGTTCTAGTACACATCTCTAAATATAGCTTTATGGCGCTTATATCTGT 3198
Db      921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
QY      3199 GATGCTTATCGCACCATCTCTGCTACTGACACACGCTCTCTATCCCATCAAGAGACATGG 3258
Db      941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY      3259 ACAACAGATCCCTTTTATAGCAAGACATGGAGTGTGTTAGAGATCTATGTATGCT 3318
Db      961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY      3319 TCCTAAAGATTAATAGAGTATATGCGCATGGAAGATATGATGATCGAGTCTCT 3378
Db      981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY      3379 CGAGGCTATGTTTGGTGCAGGAGTAGAGTCCGGTTC 3417
Db      1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

```

```

RESULT 12
US-10-498-327-3
; Sequence 3, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 02441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-3

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5121.50 Matches: 978
Percent Similarity: 98.8% Conservative: 23
Best Local Similarity: 96.5% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-3 (1-1013)
QY      382 ATGCAAACTCTTCCATAAGTTCTTCTTCAATGATTCTAGCTTATTCTGCTCTCT 441
Db      1 MetGlnThrSerPheHisLysPheLeuSerMetLeuLeuAlaTyrSerCysSer 20
QY      442 TTAATATGGGGGGATATGCAGCAGAAATCATGGTTCTCAAGGAATTTACGATGGGAG 501
Db      21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY      502 AGCTTAACGTATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTTT 561
Db      41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY      562 TCTGCAAGGAGCTTAACATTAATAAATCTTGACAAATCTTGACAGCTTTCCTTTAAGT 621
Db      61 SerAlaGlyGluLeuThrLeuLysAenLeuAenSerLeuAlaLeuProLeuSer 80
QY      622 TGTTTTGGGAACCTTATTAGGGAGTTTCTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
Db      81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY      682 GAGAAATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCCGCTGCTGATGACTG 741
Db      101 GluAenIleArgThrSerThrAenGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120
QY      742 TTTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCAATTTGCAATTCATTACTTGCCTTA 801
Db      121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140
QY      802 CTCCTGCTGCAACGACTTAATAGGGTAGCCAGACTCCGACGACCAACATCTACACCGTCT 861
Db      141 LeuProAlaAlaThrThrAenAenGlySerGlnThrProThrThrSerThrProSer 160
QY      862 AATGGTACTATTATTCTTAAACAGATCTTTTGTGTTACTCAATTAATGAGAAGTTCTCATTC 921
Db      161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAenGlyLysPheSerPhe 180
QY      922 TATAGTAATTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAGAGCTTAACCGTTCAA 981
Db      181 TyrSerAenLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY      982 GGAATTAGCAAGCTTTTGTGCTTCTTCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041

```



201 GlylleSerLysLeuCyseValPheGlnGluSerThrAlaGlnAlaAspGlyGlyAlaCys 220  
1042 CAAGTAGTCACAGTTTCTGCTATGCTTAACGAGGCTCTATTGCTTGTAGCGAAT 1101  
221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240  
1102 GTTCAGAGTAAGAGGGGAGGATTGCTGCTTTCAGGATGGGCACAGGAGTCA 1161  
241 ValAlaGlyValArgGlyGlyGlyLeuAlaValGlnAspGlyGlnGlyValSer 260  
1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACCTGCGTAGAGTTT 1221  
261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280  
1222 GATCGGAACGTAGCCCGAGTAGGAGGGATTACTCTCCAGGAACGTTGCTTTCCTG 1281  
281 AspGlyAenValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAenValAlaPheLeu 300  
1282 AATAATGAAACCTTGTCTCAACAAATGTTGCTTCTCTGTTTACATGCTGCTAAG 1341  
301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrLeuAlaGlu 320  
1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATAATTACGGAGATGGAGGACTATC 1401  
321 GlnProThrAenGlyGlnAlaSerAenThrSerAspAenTyrGlyAspGlyGlyAlaIle 340  
1402 TTCTGTAAGAAATGGTGGCAA---GCAGATCCAAATACTCGGATCAGTTTCCTTGAAT 1458  
341 PheCysLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAep 360  
1459 CGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGAAAGGGGAGCTATTAT 1518  
361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTyr 380  
1519 GCCAAAAAGCTCTCGTGTCTAACTGTGGCCCTGTACAAATTTTAAAGAAATCCCTAAT 1578  
381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400  
1579 GATGTTGAGCGGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420  
1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCAAGAGAAATGCTGCCGATGTTAAT 1698  
421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440  
1699 GCGCTAACTGTCTCAACAGCCATTTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758  
441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
1759 AGAGCTAAAGCGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818  
461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAspProIleGluMetAlaAenGlyAen 480  
1819 AACCGCCAGCCAGCTTCTCCAAACTTCTAAAAATTAACGATGGTGAAGGATACACAGG 1878  
481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAspGlyGlyLysThrGly 500  
1879 GATATTGTTTTGCTAAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAAGGA 1938  
501 AspIleValPheAlaAenGlyAenSerThrLeuTyrGlnAenValThrIleGluGlnGly 520  
1939 AGGATTGTTCTTCGTGAAAGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGTGGG 1998  
521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540  
1999 ACTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACCAACCAACAG 2058  
541 SerLeuTyrMetGluAlaGlySerThrLeuAepPheValThrProGlnProGlnGln 560  
2059 CCTCTGCCCGCTAAATCAGTTGATCAGCTTTTCCAATCTGCATTTGTCTCTTCTTTG 2118

561 ProProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeu 580  
2119 TTACCAACAATGAGTTACGAATCCTCTACCAATCCTCCAGCGCAGAGATTCATCATCT 2178  
581 LeuAlaAenAenAlaValThrAenProThrAenProAlaGlnAspSerHisPro 600  
2179 GCAGTCTTGTGTAGCACAACTGCTGTTCTTCAATATTAGTGGGCCCTATCTTTTTTCAG 2238  
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
2239 GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGTCTTAATCAAAAATCAAT 2298  
621 AspLeuAepAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAenGlnLysIleAep 640  
2299 GTCTGAAATTTACAGTTAGGAGCTAAGCCCCACCTAATGCCCATCAGATTTGACTCTA 2358  
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660  
2359 GGGATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGCCTGGATCCTAAT 2418  
661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaIlePhePheProAen 680  
2419 ACAGCAATATGCTCTTATCTACTCTGAAAGCTACATGAGCTAAACCTGGGTATAATCCT 2478  
681 ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAenPro 700  
2479 GGGCTGAGCAGTAGCTTCTTGGTTCCAAATAGTTTATGGGATCCATTTAGATATA 2538  
701 GlyProGluArgValAlaSerLeuValProAenSerLeuTyrGlySerIleLeuAspIle 720  
2539 CGATCTGGCATTCAGCAATTCAGAGCAAGTGTGATGGCGCTCTTATTGTTCGAGATTA 2598  
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyLysSerTyrCysArgGlyLeu 740  
2599 TGGTGTCTGAGATTTCGAATTTCTCTATCATGACCCGATGCTTTTAGTTCAGGAGAT 2658  
741 TrpValSerGlyValSerAenPheTyrHisAepArgAepAlaLeuGlyGlnGlyTyr 760  
2659 CGGTATATTAGTGGGGTTATTTCTTAGGACCAACTCTCTACTTTGGATCATCGATGTTT 2718  
761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780  
2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATATTAGTGTGCTGTTCCCAAT 2778  
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800  
2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCCCAACAGCTTTTATGATGATCCAT 2838  
801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
2839 TTGTTCCGAGATGCGTTTATCCGTGCTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898  
821 LeuPheGlyAepAlaPheLeuArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840  
2899 TCATATACATTTGCAGAGGAGCGATGCTCGTTGGGATAATACTGCTGCTGCGAGAG 2958  
841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAenAenCysLeuValGlyGlu 860  
2959 ATTGAGCGGGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGAAATGAGTTGCGT 3018  
861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880  
3019 CCTTTGCGAAGCTGAGTTTCTTATCCGATCATGATCTTTTACAGAGGAGCGGAT 3078  
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAep 900  
3079 CAAGCTCCGGCATTCAGAGGAGCATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138  
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920  
3139 TTTGATCGATGTTCTAGTACACATCCTTAATAATAGCTTTATGGCGGCTTATCTGT 3198  
921 PheAepArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940



QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGCAACGCTCTATCCATCCATCAAGAGCATGG 3258  
DB 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCGCTTTTCATTTAGCAAGACATGGAGTTGGTTAGAGGATCTATGATGCT 3318  
DB 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValAlaGlySerMetTyrAla 980  
QY 3319 TCTTAAACAGTAATATAGAAATATATGCGCATGGAGATATGAGATATGAGATGCTTCT 3378  
DB 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
QY 3379 CGAGGCTATGTTTGAAGTCAGGACAGTACAGTCCGGTTC 3417  
DB 1001 ArgGlyTyrGlyLeuSerAlaGlySerIleValAlaGlyPhe 1013

RESULT 13  
US-10-701-844-43  
; Sequence 43, Application US/10701844  
; Publication No. US20040067524A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/10/701,844  
; CURRENT FILING DATE: 2003-11-04  
; PRIOR APPLICATION NUMBER: US/09/612,402B  
; PRIOR FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 43  
; LENGTH: 984  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-701-844-43

Alignment Scores:  
Pred. No.: 0 Length: 984  
Score: 5116.00 Matches: 984  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 64.9% Indels: 0  
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-701-844-43 (1-984)

QY 466 GAAATCATGTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGATCATTTCCCTAT 525  
DB 1 GluIleMetValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20  
QY 526 ACTGTTATAGAGATCCGATGGGACTGCTTTTCTCAGGAGAGTTAAACATTAAGA 595  
DB 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40  
QY 586 AATCTTGACAAATCTATTGACGCTTTCCTTTTAAAGTTGTTTGGAACTTTATTAGGGAGT 645  
DB 41 AsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60  
QY 646 TTTACTGTTTTAGGAGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACAAT 705  
DB 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80  
QY 706 GGGGCGACTTAAGTAATAGCGTCTGCTGAGGACTGTTTACTATTGAGGGTTTTAAGAA 765  
DB 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100  
QY 766 TTATCTTTTCCAAATTCATTTACTTACTTGGCGTACTGCTGCTGCAACGACTTAATAAG 825  
DB 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120

QY 826 GGTAGCAGACTCCGACGACACATCTACACCGTCTAATGGTACTATTATTCTAAACA 885  
DB 121 GlySerGlnThrProThrThrThrSerProSerAsnGlyThrIleTyrSerLysThr 140  
QY 886 GATCTTTTGTACTCAATAAGTCAAGAGTCTCATTTCTATAGTAATTTAGTCTCTCGAGAT 945  
DB 141 AspLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAsp 160  
QY 946 GGGGCGAGCTATAGATCTTAAGAGCTTAACGGTTCAAGGAATTTAGCAAGCTTTGTCTTC 1005  
DB 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180  
QY 1006 CAAGAAAAATCTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAAGTTTCTCTGCT 1065  
DB 181 GlnGluAsnThrAlaGlnAlaAspGlyAlaCysGlnValValThrSerPheSerAla 200  
QY 1066 ATGGCTAAGAGGCTCTTATTCCTTTGTAGCGAATGTTGCAGGAGTAGAGGGGAGGG 1125  
DB 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220  
QY 1126 ATTGCTGCTGTTTCAGGATGGGCGAGGAGTGTATCATCTACTTCAACAGAGATCCA 1185  
DB 221 IleAlaAlaValGlnAspGlyGlnGlyValSerSerThrThrSerThrGluAspPro 240  
QY 1186 GTAGTAAGTTTTCAGAAATACTCGGCTAGAGTTTGTATGGGAACGTAGCCGAGTAGGA 1245  
DB 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260  
QY 1246 GAGGAGATTTACTCTTACCGGAACTGCTTCTCTGCTCAATATGGAACCTTGTCTTC 1305  
DB 261 GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280  
QY 1306 AACATGTTGCTTCTCTGCTTACATGCTGTCTTAAGCAACCAACAGTAGTGACAGGCTTCT 1365  
DB 281 AsnAsnValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300  
QY 1366 AATACGAGTAATAATTTACCGAGATGGAGGAGCTATCTTCTGTAAAGAAATGGTGCAGCA 1425  
DB 301 AsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320  
QY 1426 GGATCCAATACTCTGGATCAGTTCTCTTGTATGGAGAGGAGTAGTTTCTTTTAGTAGC 1485  
DB 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340  
QY 1486 AATGATGCTCTGGGAAAGGGGAGCTATTATGCAAAAAGCTCTCGGTGCTAACTGT 1545  
DB 341 AsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCys 360  
QY 1546 GGCCCTGTACAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGATTTATTATTAGGAA 1605  
DB 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380  
QY 1606 TCTGGAGAGCTCAGTTTATCTGCTGATATGGAGATATTATTTTCAGTGGGAATCTTAAA 1665  
DB 381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400  
QY 1666 AGAACGCCAAGAGAAATGCTGCCGATGTTAATGGCGTAATCTGCTCTCACAGCCATT 1725  
DB 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420  
QY 1726 TCCATGGATCGGGAGGAAATAACGACATTAAAGAGCTAAAGCAGGCGCATCAGATTCTC 1785  
DB 421 SerMetGlySerGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440  
QY 1786 TTTAATGATCCCATCGAGATGGCAACGGAAATAACGACGAGCGAGTCTTCCAAACTT 1845  
DB 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460  
QY 1846 CTTAAATTAACGATGTTGAGGATACACAGGGGATATCTGTTTGTGATGAGAGCAGT 1905  
DB 461 LeuLysIleAsnAspGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480

```
QY 1906 ACTTTGTACAAAATGTTACGATAGACGAAAGGATTGTTCTTCTGTGAAAAGCGCAAAA 1965
Db ThrLeuTyrGlnAenValThrIleGluGlnGlyArgIleValLeuArgGluLeuAalalys 500
QY 1966 TTATCAGTGAATTCCTTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGCAGTACA 2025
Db LeuSerValAenSerLeuSerGlnThrGlySerLeuTyrMetGluAalaglySerThr 520
QY 2026 TGGGATTTGTAACTCCACAACCCACCAACAGCCTCCTGCCGCTTAATCAGTTGATCACG 2085
Db TrpAspPheValThrProGlnProProGlnProProAlaAlaAenGlnLeuIleThr 540
QY 2086 CTTTCCAAATCTGCATTTGCTCTTTCTTTCTTTGTTAGCAACAATGCAGTTACGAATCCT 2145
Db LeuSerAenLeuHleLeuSerLeuSerSerLeuLeuAlaAenAenAlaValThrAenPro 560
QY 2146 CCTACCAATCTCCAGCGCAAGATTCTCATCTCGCAGTCATTGGTAGCACAACTGCTGGT 2205
Db ProThrAenProProAlaGlnAspSerHisProAlaValIleGlySerThrAlaGly 580
QY 2206 TCTGTTACAAATTAGTGGCCTATCTTTTTCAGGATTTGGATGATACAGCTTATCATAGG 2265
Db SerValThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArg 600
QY 2266 TATGATTCGGTAGGTTCTTAATCAAAAATCAATGTCCTGAAATTTACAGTTAGGAGCTAAG 2325
Db TyrAspTrpLeuGlySerAenGlnIlyIleAenValLeuLysLeuGlnLeuGlyThrLys 620
QY 2326 CCCCAGCTAATGCCCATCAGATTGACTCTAGGGAATGAGTCCCTAAGTAGGCTAT 2385
Db ProProAlaAenAlaProSerAspLeuThrLeuGlyAenGluMetProLysTyrGlyTyr 640
QY 2386 CAAGGAAGCTGGAAGCTTCGCTGGGATCTTAATACAGCAAAATAATGGTCCTTATCTCTG 2445
Db GlnGlySerTrpLysLeuAlaTrpAspProAenThrAlaAenAenGlyProTyrThrLeu 660
QY 2446 AAAGCTACATGAGCTAAAGCTGGGTATTAATCTCGGCCCTGAGCGAGTAGCTTCTTTGGTT 2505
Db LysAlaThrTrpThrLysThrGlyTyrAenProGlyProGluArgValAlaSerLeuVal 680
QY 2506 CCNAATAGTTATGGGGATCCATTTTATAGATATACGATCTGGCATTCAGCAATTCAGCA 2565
Db ProAenSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAla 700
QY 2566 AGTGTGATGGCGCTCTTATTTGTCGAGATTTATGGTTCTTCGAGTTTCGAAATTCCTC 2625
Db SerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAenPhePhe 720
QY 2626 TATCATGACCGGATGCTTTAGTTCAGGATATCGGATATATAGTGGGGTTATTCCTTA 2685
Db TyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyTyrSerLeu 740
QY 2686 GGAGCAAACTCTACTTTGGATCATCGATGTTTGGTCTAGCATTTACCGAAGTATTTGCT 2745
Db GlyAlaAenSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGly 760
QY 2746 AGATCTAAAGATTATGTAGTGTGCTTCCCAATCATCATGCTTGCATAGGATCCGCTTAT 2805
Db ArgSerLysAspTyrValValCysArgSerAenHisHisAlaCysIleGlySerValTyr 780
QY 2806 CTATCTACCCAAACAGCTTTATGTGATCTTATTTGTTTCGGAGATGCGTTTATCCGTGCT 2865
Db LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAla 800
QY 2866 AGCTACGGTTTGGCAATCAGCATATGAAAACCTCATATACATTTTCGACAGGAGCGCAT 2925
Db SerTyrGlyPheGlyAenGlnHisMetLysThrSerTyrThrPheAlaGlnGluSerAsp 820
QY 2926 GTTCGTGGGATAAATAACTGCTGCTGAGAGATTGGACGGGATTTACCGATTGTGATT 2985
Db ValArgTrpAspAenAenCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840
QY 2986 ACTCCATCTAAGCTCTATTGTAATGAGTTGCGTCTCTTCGTGCAAGCTGAGTTCTTAT 3045
```

```
Db 841 ThrProSerLysLeuTyrLeuAenGluLeuArgProPheValGlnAlaGluPheSerTyr 860
QY 3046 GCCGATCATGAATCTTTTACAGAGGAAGGCGATCAAGCTCGGCGATTTCAAGAGCGGACAT 3105
Db AlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHis 880
QY 3106 CTCCTAAATCTATCAGTTCCTGTTGGAGTAGGTTGATCGATGTTCTTAGTACACATCCT 3165
Db LeuLeuAenLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisPro 900
QY 3166 AATAAATATAGCTTTATGGCGCTTATATCTGTGATCTTATCGCACCATCTCTGGTACT 3225
Db AsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThr 920
QY 3226 GAGCAACAGCTCTCCATCCCATCAAGACATGGAACAACAGATGCTTTTCATTTAGCAAGA 3285
Db GluThrThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArg 940
QY 3286 CATGGAGTTGTTGTTAGAGATCTATGTATGCTTCTTAACAAGTAATATAGAGTATAT 3345
Db HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAenIleGluValTyr 960
QY 3346 GGCCATCGAAGATATGATATCGAGATGCTTCTCGAGCTATGTTGAGTGCAGGAAGT 3405
Db GlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980
QY 3406 AGAGTCCGGTTC 3417
Db ArgValArgPhe 984

RESULT 14
US-10-766-711-43
; Sequence 43, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10766,711
; PRIORITY FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-43

Alignment Scores:
Pred. No.: 0 Length: 984
Score: 5116.00 Matches: 984
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.9% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-766-711-43 (1-984)
QY 466 GAAATCATGTTCTCTCAAGGAATTTACGATGGGAGAGCTTAACGTGATCATTTCCCTAT 525
Db 1 GluIleMetValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20
QY 526 ACTGTTATAGAGATCCGAGTGGGACTACTGTTTTTTCTCGAGGAGCTTAACATTAAA 585
Db 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
QY 586 AATCTTGACAAATCTTATTGACAGCTTTGCTTTTAAAGTTGTTTGGGAACCTTATTAGGAGT 645
Db 41 AsnLeuAspAenSerIleAlaLeuProLeuSerCysPheGlyAenLeuLeuGlySer 60
```

QY 646 TTTACTGTTTGGGAGGACACCTCGTTGACTTTCGAGACATACGAGCTTCTACAAT 705  
 Db PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80  
 QY 706 GGGGAGCTCTAAGTAATAGCCCTGCTGATGACTGTTTACTATTATGAGGGTTTAAAGAA 765  
 Db 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheGlyGlu 100  
 QY 766 TTATCCTTTTCCAATTTGCAATTCATTACTGTCGGTACTGCTCGTACGACGCTAATAG 825  
 Db 101 LeuSerPheSerAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120  
 QY 826 GGTAGCCAGACTCCGACGACACACTACACCGCTCTAATGGTACTATTATTCTAAACA 885  
 Db 121 GlySerGlnThrProThrThrThrThrProSerAsnGlyThrIleIleThrSerLysThr 140  
 QY 886 GATCTTTTGTACTCAATTAATGAGAAGTTCTCATCTATAGTAATTTAGTCTCTGAGAT 945  
 Db 141 AspLeuLeuLeuLeuAsnAsnGluLysPheSerPheThrSerAsnLeuValSerGlyAsp 160  
 QY 946 GGGGAGCTATAGACTTAAGACTTAACGGTTCAAGGAATTAGCAGCTTTGTGCTCTC 1005  
 Db 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180  
 QY 1006 CAAGAAATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACACCACTTCTCTGCT 1065  
 Db 181 GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAla 200  
 QY 1066 ATGGCTTAACGAGCTCTATTGCTTTGTAGCAATGTTGCAGGAGTAAGAGGGGAGGG 1125  
 Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValargGlyGlyGly 220  
 QY 1126 ATTGCTGCTTTCAGGATGGGAGGAGGTGTCACTATCTACTTCAACAGAGATCCA 1185  
 Db 221 IleAlaAlaValGlnAspGlyGlnGlyValSerSerSerThrSerThrGluAspPro 240  
 QY 1186 GTAGTAAGTTTTCCAGAAATACTGCGGTAGAGTTTGTGGGACCTAGCCGAGTAGGA 1245  
 Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260  
 QY 1246 GAGGGATTTACTCTACGGGACGTTGCTTTCTCGAATAATGAAAAACCTTTGTTCTC 1305  
 Db 261 GlyGlyIleThrSerThrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280  
 QY 1306 AACCAATGTTGCTTCTCTGTTTACATTGCTGTGCTPAAGCAACCAACAGTGACAGGCTTCT 1365  
 Db 281 AsnAsnValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300  
 QY 1366 AATACGAGTAATTAATTCGAGATGGAGGAGCTATCTTCTGTAAGATGGTGGCGAGCA 1425  
 Db 301 AsnThrSerAsnAsnThrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320  
 QY 1426 GGATCCAATAACTCTCGATCAGTTTCTTTGATGGAGAGGAGTAGTTTTCTTTAGTAGC 1485  
 Db 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSer 340  
 QY 1486 AATGTAGCTCTCGGAAAGGGGAGCTATTATTATGCCAAAAGCTCTCGGTGCTAACTGT 1545  
 Db 341 AsnValAlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCys 360  
 QY 1546 GGCCTGTACAATTTTAAAGAAATATCGCTAATGATGGTGGAGGATTTATTATTAGGAA 1605  
 Db 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380  
 QY 1606 TCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTTCCAGTGGATCTTAAA 1665  
 Db 381 SerGlyGlnLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400  
 QY 1666 AGAACACCCAAAGAGATGCTGCGATGTTAATGGCGTAACTGTGTCTCTCACAGCCATT 1725  
 Db 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420

QY 1726 TCGATGGATCGGAGGAGGAAAAATAACGATTAAGAGCTAAAGCAGCGCATCAGATTCTC 1785  
 Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440  
 QY 1786 TTTAATGATCCCATCGAGATGCAACCGAAATAACAGCCAGCGCAGCTCTCCAAACTT 1845  
 Db 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460  
 QY 1846 CTAATAATTAACGATGGTGAAGGATACACAGGGGATATTGTTTTGCTTAATCGAAGCAGT 1905  
 Db 461 LeuLysIleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480  
 QY 1906 ACTTTGTACAAAATTTTACGATAGCAAGGAAGGATTTCTTCCTCGTGAAGGCAAAA 1965  
 Db 481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500  
 QY 1966 TTATCAGTGAATTTCTTAAGTCAGAGCTGGAGTCTGTATATGGAAGCTCGGAGTACA 2025  
 Db 501 LeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThr 520  
 QY 2026 TGGGATTTTGTAACTCCACCAACCAACCAACCAACCTCTGCGCGCTAATCAGTTGATCAGC 2085  
 Db 521 TrpAspPheValThrProGlnProGlnProGlnProAlaAlaAsnGlnLeuIleThr 540  
 QY 2086 CTTTCCAATCTGCATTTGCTCTTTCTTTGTAGCAAAACAATGCGAGTTAGGATCCT 2145  
 Db 541 LeuSerAsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAlaValThrAsnPro 560  
 QY 2146 CTTACCAATCTCGAGCGCAAGATTCTCATCTGCTGAGTCATTTGTTAGCAACAATCGCTGT 2205  
 Db 561 ProThrAsnProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGly 580  
 QY 2206 TCTGTTTACAATTAGTGGGCTATCTTTTGTAGGATTTTGCATGATACAGCTTATGATAGG 2265  
 Db 581 SerValThrIleSerGlyProIlePhePheGluAspLysAspThrAlaTyrAspArg 600  
 QY 2266 TATGATTTGGCTAGGTTCTAATCAAAAAATCAATGTCTCGAAATACAGTTAGGAGCTAAG 2325  
 Db 601 TyrAspTrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLys 620  
 QY 2326 CCCCAGCTAATGCCCCCATCAGATTGACTCTAGGGAATGAGATGCTAAGTATGCTAT 2385  
 Db 621 ProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyr 640  
 QY 2386 CAAGGAGCTCGAAGCTCTGCTGGATCTTAATACAGCAAAATAATGCTCTTATATCTCTG 2445  
 Db 641 GlnGlySerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeu 660  
 QY 2446 AAAGCTTACATGGAATAAACTGGGTATAATCTCGGCTGAGCGAGTAGCTCTTTGGTT 2505  
 Db 661 LysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuVal 680  
 QY 2506 CCAATAGTTTATGGGATCCATTTTAGATATACGATCTCGCATTCAGCAATTCAGCA 2565  
 Db 681 ProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAla 700  
 QY 2566 AGTGTGGATGGGCTCTTATTGTGAGGATATGCTGCTTCTGAGTTTCGAATTTCTTCT 2625  
 Db 701 SerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePhe 720  
 QY 2626 TATCATGACCGCATCTTTTAGTTCAGGATATCGGTATATTTAGTGGGGTATTCTTTA 2685  
 Db 721 TyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeu 740  
 QY 2686 GGAGCAAACTCTTACTTTTGGATCATCGATTTTGGTCTAGCATTTACCGAAGTATTGGT 2745  
 Db 741 GlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGly 760  
 QY 2746 AGATCTAAAGATTATGTAGTGTCTGCTTCCAAATCATCATCTGCTTAGGATCCGTTTAT 2805  
 Db 761 ArgSerLysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyr 780  
 QY 2806 CTATCTACCCCAACAGCTTTTATGTGGATCTTATTTGTTGCGAGATGCTTTATCCGTGCT 2865

```
Db      781  LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleAargAla 800
Qy      2866 AGCTACGGGTTGGGAATCAGCATATGAAACCTCATATACATTTGCAGAGGAGCGCAT 2925
Db      801  SerTyrGlyPheGlyAsnGlnHisMetIlySerTyrThrPheAlaGluGluSerAsp 820
Qy      2926 GTTCGTTGGGATAATACTGCTGGCTGGAGAGATTGGAGCGGGAATACCGAATTGTGATT 2985
Db      821  ValArgTyrAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840
Qy      2986 ACTCCATCAAGCTTATTGATGAGTGGCTGCTTTCGTCGAGCGAGTCAGTCTCTTAT 3045
Db      841  ThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyr 860
Qy      3046 GCCGATCATGAATCTTTTACAGAGAGCGCATCAAGCTCGGGCATTCAGAGCGGACAT 3105
Db      861  AlaPheHisGluSerPheThrGluGlyAspGlnAlaAargAlaPheLysSerGlyHis 880
Qy      3106 CTCCTAAATCATCAAGTTCCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCT 3165
Db      881  LeuLeuAsnLeuSerValProValGlyValLysPheAspAargCysSerSerThrHisPro 900
Qy      3166 AATAAATATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTACT 3225
Db      901  AsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThr 920
Qy      3226 GAGACACGCTCTATCCATCCATCAAGAGACATGACACAGATGCTCTTCATTAGCAAGA 3285
Db      921  GluThrThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaAarg 940
Qy      3286 CATCGAGTTGTGGTATGAGGATCTATGATGCTTCTTAACAAGTAATATAGAATATAT 3345
Db      941  HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyr 960
Qy      3346 GCCCATGAGATATGATATCGATGCTTCTCGAGGCTATGCTTGTGATGTCAGGAAGT 3405
Db      961  GlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980
Qy      3406 AGAGTCCGGTTC 3417
Db      981  ArgValArgPhe 984
```

## RESULT 15

```
US-10-931-779-43
; Sequence 43, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-43
```

```
Alignment Scores:
Pred. No.: 0 Length: 984
Score: 5116.00 Matches: 984
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.9% Indels: 0
DB: Gaps: 0
```

US-10-701-844-1 (1-4435) x US-10-931-779-43 (1-984)

```
Qy      466  GAAATCATGTTCTCTCAAGGAATTTACATGGGAGAGCTTAACTGTATCATTTCCCTAT 525
Db      1  GluIleMetValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20
Qy      526  ACTGTTATAGAGATCCGAGTGGGACTACTGTTTTTCTCGCAGGAGAGCTTAACATTAAAA 585
Db      21  ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
Qy      586  AATCTTGACAATTTCTATTGCAAGCTTTGCTTTAAAGTTGTTTTGGAACTTATTAGGAGT 645
Db      41  AsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60
Qy      646  TTTACTGTTTGGGAGAGGACACTCGTTGACTTTCGAGAACATACGAGCTTCTACAAT 705
Db      61  PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
Qy      706  GGGCGAGCTCTAAGTAAATAGCGCTGATGAGCTGTTTTACTATTAGGGTTTTTAAAGAA 765
Db      81  GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100
Qy      766  TTATCTCTTTTCCAATTTGCAATTTCAATTTCTGCCGTACTGCTGCTGCAACGACTTAATAAG 825
Db      101  LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
Qy      826  GGTAGCCAGACTCCGACGACCAATCTACACCGCTCTAATGCTACTATTATTTCTTAAACA 885
Db      121  GlySerGlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThr 140
Qy      886  GATCTTTTGTACTCAATAATGAGAGTTCTCATTCTATAGTAAATTTAGTCTCTGAGAT 945
Db      141  AspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAsp 160
Qy      946  GGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTCTTC 1005
Db      161  GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
Qy      1006  CAAGAAATATCTGCTCAAGCTGATGGGAGCTGTTCAAGTAGTACACAGTTTCTCTGCT 1065
Db      181  GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAla 200
Qy      1066  ATGGCTAACGAGGCTCTATTGCTTCTAGCAATGTTGAGGAGTAGTAAGGGGAGGG 1125
Db      201  MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
Qy      1126  ATTGCTGCTGTTTCAGGATGGGAGGAGGTGTCATCATCTACTTCAACAGAAATCCA 1185
Db      221  IleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrSerThrGluAspPro 240
Qy      1186  GTAGTAAGTTTTTCCAGAAATACCTCGGTAGAGTTTGTGGAACGTAGCCCGAGTAGGA 1245
Db      241  ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
Qy      1246  GGAGGGATTTACTCTCTAGGGAACGTGCTTCTCGTAATATGGAACCTTGTCTCTC 1305
Db      261  GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280
Qy      1306  AACCAATGTTGCTTCTCTGTTTACATGCTGCTTAAGCAACCAACCAAGTGGACAGGCTCT 1365
Db      281  AsnAsnValAlaSerProValTyrIleAlaLysGlnProThrSerGlyGlnAlaSer 300
Qy      1366  AATACGAGTAATAATTACGAGATGGAGAGCTATCTTCTGTAAAGATGGTGGCAGCA 1425
Db      301  AsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320
Qy      1426  GGATCCAATACTCTGGATTCAGTTTCTTGTATGAGAGGGAGTAGTTTCTTTAGTAGC 1485
Db      321  GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340
Qy      1486  AATGTAGCTGCTGGGAAAGGGGAGCTATTATATGCAAAAAGCTCTCGGTCTCACTGT 1545
Db      341  AsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCys 360
```

1546 GGCCTGTACAAATTTTAAAGCAATATCGCTAATGATGTCGAGCGATTATTTAGGAGAA 1605  
Db |||||  
361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380  
Qy |||||  
1606 TCTGAGAGCTCAGTTTATCTGCTGATATGAGATATATTTTCGATGGGAATCTTAAA 1665  
Db |||||  
381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400  
Qy |||||  
1666 AGACAGCCAAAGCAAGTCTGCCGATGCTTAATGCGCTAATGTCCTCACAGCCATT 1725  
Db |||||  
401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420  
Qy |||||  
1726 TCGATGGATCGGAGCGGAAATTAACAGCTTAAGAGCTAAAGCAGGCGCATCTCTC 1785  
Db |||||  
421 SerMetGlySerGlyGlyIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440  
Qy |||||  
1786 TTTAATGATCCCATCGAGATGGCAACCGAAATAACAGCCAGCGCAGCTCTTCCAACTT 1845  
Db |||||  
441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460  
Qy |||||  
1846 CTAAATAATTACGATGGTAGGATACACAGGGATATGTTTTGCTTAATGGAGCAGT 1905  
Db |||||  
461 LeuLysIleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480  
Qy |||||  
1906 ACTTTGTACCAAAATGTTACGATAGCAGCAAGGAGATTGTTCTTCGTGAAAAGCAAAA 1965  
Db |||||  
481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500  
Qy |||||  
1966 TTATCAGTGAATTTCTTAAGTCAGACAGTGGGAGTCTGTATATGAAAGCTGGGAGTACA 2025  
Db |||||  
501 LeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThr 520  
Qy |||||  
2026 TGGGATTTGTAACTCCCAACCAACCAACAGCTCTCGCGCTAATCAGTTGATCAGC 2085  
Db |||||  
521 TrpAspPheValThrProGlnProGlnGlnProAlaAlaAsnGlnLeuIleThr 540  
Qy |||||  
2086 CTTTCCAAATCTCATTGCTCTTCTTCTTCTTGTAGCAAAACAATGCAATCTACGAATCT 2145  
Db |||||  
541 LeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAlaValThrAsnPro 560  
Qy |||||  
2146 CCTACCAATCTCCAGCCAGATTTCTCATCTCGCAGTCATTTGGTAGCAAACTCGTGT 2205  
Db |||||  
561 ProThrAsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGly 580  
Qy |||||  
2206 TCTGTTACAATTAGTGGGCTATCTTTTGTAGGATTTGGATGATACAGCTTATGATGG 2265  
Db |||||  
581 SerValThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArg 600  
Qy |||||  
2266 TATGATTCGCTAGGTCTTAATCAAAAAATCAATGCTCGAAATTAACAGTTAGGACTAAG 2325  
Db |||||  
601 TyrAspTrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLys 620  
Qy |||||  
2326 CCCCCAGCTTAATGCCCCATCAGATTTGACTCTAGGGAATGAGATGCTTAAGTATGCTAT 2385  
Db |||||  
621 ProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyr 640  
Qy |||||  
2386 CAAGGAAGCTGAAGCTTCGCTGGGATCTTAATACAGCAAAATAATGGTCTTATATCTG 2445  
Db |||||  
641 GlnGlySerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeu 660  
Qy |||||  
2446 AAGCTTACATGACCTAAACTGGGTATATCTCGGCGCTGAGCAGTACTCTTTGGTT 2505  
Db |||||  
661 LysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuVal 680  
Qy |||||  
2506 CCAATATAGTTTATGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCA 2565  
Db |||||  
681 ProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAla 700  
Qy |||||  
2566 AGTGTGATGGGCGCTCTTATTTGTGAGGATTAATGGTTCTGAGTTTCGAATTTCTTC 2625  
Db |||||  
701 SerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePhe 720  
Qy |||||  
2626 TATCATGACCGCATCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATTCCTTA 2685

721 TyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeu 740  
Qy |||||  
2686 GGACCAACTCCTACTTTGGATCATCGATGTTTGGTCTTAGCATTTACCGAAGTATTTCGT 2745  
Db |||||  
741 GlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGly 760  
Qy |||||  
2746 AGATCTAAAGATTTAGTGTGTCTGTTCCAAATCATCATGCTTCATAGGATCCGTTTAT 2805  
Db |||||  
761 ArgSerLysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyr 780  
Qy |||||  
2806 CTATCTACCCAAAGCTTTATGTCGATCCTTATTTGTCGAGATGCGTTTATCCGTCCT 2865  
Db |||||  
781 LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAla 800  
Qy |||||  
2866 AGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTTGCAGAGAGAGCCAT 2925  
Db |||||  
801 SerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAsp 820  
Qy |||||  
2926 GTTCGTTGGGATAATAACTGTCTGCTGGAGAGATTGGAGCGGATTAACCGATTGTGATT 2985  
Db |||||  
821 ValArgTrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840  
Qy |||||  
2986 ACTCCATCTAAGCTCTATTTCGAATGAGTTGCGTCTCTTTCTGTCGAAAGCTGAGTTTCTTAT 3045  
Db |||||  
841 ThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyr 860  
Qy |||||  
3046 CCGCATCATGATCTTTTACAGAGCAAGCGCATCAAGCTCGGGCATTCGAAGCGGACAT 3105  
Db |||||  
861 AlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHis 880  
Qy |||||  
3106 CTCCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTTGTATCGATGTTCTAGTACACATCT 3165  
Db |||||  
881 LeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisPro 900  
Qy |||||  
3166 AATAAATATAGCTTTATCGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTACT 3225  
Db |||||  
901 AsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThr 920  
Qy |||||  
3226 GAGCAACGCTCTTATCCCATCMAGACATGGAACAACAGATGCTCTTTCATTTAGCAAGA 3285  
Db |||||  
921 GluThrThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArg 940  
Qy |||||  
3286 CATGAGTGTGTTAGAGGATCTATGATGCTCTCTAAACAAGTAATATAGAAGTATAT 3345  
Db |||||  
941 HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyr 960  
Qy |||||  
3346 GGCCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGGTTTGTGTCAGGAAGT 3405  
Db |||||  
961 GlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980  
Qy |||||  
3406 AGAGTCCGGTTC 3417  
Db |||||  
981 ArgValArgPhe 984

RESULT 16  
US-10-498-327-17  
; Sequence 17, Application US/10498327  
; Publication NO. US20050106162A1  
; GENERAL INFORMATION:  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis  
; FILE REFERENCE: 002441.00085  
; CURRENT APPLICATION NUMBER: US/10/498,327  
; CURRENT FILING DATE: 2004-06-10  
; PRIOR APPLICATION NUMBER: PCT/IB02/05761  
; PRIOR FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 262  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 17  
; LENGTH: 1013  
; TYPE: PRT

ORGANISM: Chlamydia trachomatis  
US-10-498-327-17

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5112.50 Matches: 977  
Percent Similarity: 98.7% Conservative: 23  
Best Local Similarity: 96.4% Mismatches: 12  
Query Match: 64.9% Indels: 1  
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-17 (1-1013)

```
QY 382 ATGCAACGCTTTCCATAAGTTCTTTCTTCAATGATTTAGCTATTCTTGCTGCTCT 441
DB 1 MetClnThrSerPheHisLysPhePheLeuSerMetLeuLeuAlaTyrSerCysSer 20
QY 442 TTAATGGGGGGATATGACGACGAGAAATCATGTTCTCTCAAGAAATTTACGATGGGAG 501
DB 21 LeuSerGlyGlyGlyTyrAlaAlaGluMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTGCTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGATTAACATTAATAAAATCTTCACAAATCTATTGACGCTTTGCCCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80
QY 622 TGTGTTGGGAATTTATAGGAGTTTTACTGTTTTAGGAGAGGACACTCGTTCACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGACTTCTCAAAATGGGGCAGCTCTAAGTAATAGCGTGTGATGGGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTACTATTGAGGTTTTAAAGATTATCTTTTCCATTCAATTCATTACTTGGCGTA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnGlyAsnSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTTAATAAGGTAGCCAGACTCCGACGACAACTATCACCGTCT 861
DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGTTACTATTATCTCAAAACAGACTCTTTGTTTACTCAATATGAGAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAGCTTTGTGTTCTTCAAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGlnLeuThrAlaGlnAlaAspGlyAlaCys 220
QY 1042 CAAGTAGTACACAGTTTCTGCTATGCTTAAGCTTAAGAGCTCTTATGCTTTGCTAGGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
QY 1102 GTTGCAGGAGTAAGGGGGAGGATGCTGCTGTTTCAGATGGGACGAGGAGTGCA 1161
DB 241 ValAlaGlyValArgGlyGlyLeuAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATATCTCCGCTAGAGTTT 1221
DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACGTACCGGAGTAGGAGGAGTTTACTCTACGAGGAACTGCTCTTCCTG 1281
DB 281 AspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
```

```
QY 1282 AATAATGGAAAAACCTTGTTTCTCAACAATGTTGCTTCTCCTGTTTACATTGCTGCTAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrLeuAlaAlaGlu 320
QY 1342 CAACCAACAGTCGACAGGCTTCTAATACGAGTAATAATTTACGAGATGAGGAGCTATC 1401
DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAsnTyrGlyAspGlyGlyAlaIle 340
QY 1402 TTCTGTAAGAAATGCTGCGCAA---GCAGGATCCAATAACTCTCGATCAGTTCTCTTGAT 1458
DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 GGAGAGGAGTAGTTTCTTTTAGTCAATGTAGTCTGCTGGGAAAGGGGAGCTATTAT 1518
DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380
QY 1519 GCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578
DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
QY 1579 GATGCTGAGCGATTTATTTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638
DB 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTATTTTGAATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAT 1698
DB 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GGCGTAATCTGTCCTCACAAAGCCATTTCCGATGGGATGGGAGGAAAAATAACACATTA 1758
DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyPheLysIleThrThrLeu 460
QY 1759 AGACTAAAGCAGGCGCATCAGATTCTTTAATGATCCCATCGAGATGGCAACCGAAAT 1818
DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACGAGCCAGCGAGTCTTCCAACTTCTAAAATTAAACGATGTGTAAGGATACACAGGG 1878
DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyrThrGly 500
QY 1879 GATATTGTTTGTCTTAATGAGCAGTACTTTGACCAAAATGTTTACGATAGACAGGA 1938
DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATTGTTCTTCGTAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGG 1998
DB 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCACACAG 2058
DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTCCATCTGCAATTTGCTCTTCTTCTTTG 2118
DB 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAACAATGCAAGTTACGAATCTCTACCAATCTCCAGCGCAAGATTCATCCT 2178
DB 581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCATGTGTAGCACAACTGCTGTTCTGTATCAATTTAGTGGGCTATCTTTTTCAG 2238
DB 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAAATCAAAAATCAAT 2298
DB 621 AspLeuAspAspThrAlaTyrAspArgTyrAspIlePleuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCCGAAAATACAGTTAGGGACTAAGCCGCCAGCTAATGCCCCATCAGATTGACTCTA 2358
DB 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGAAATGAGTCCCTAAAGTATGGCTATCAAGGAGCTGGAAGCTTGGGGATCCTAAT 2418
```



```

Db      661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAen 680
QY      2419 ACAGCAAAATAATGGTCCCTTATCTCTGAAAGCTACATGAGCTAAACATGGGTATAATCTCT 2478
Db      681 ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAenPro 700
QY      2479 GGGCTGAGCGAGTAGCTCTTGGTGTCCAAATAGTTTATGGGATCCATTTTATAGATATA 2538
Db      701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
QY      2539 CCATCTGCCCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTGCGAGATTA 2598
Db      721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY      2599 TGGGTTCTGGAGTTTCGAATTTCTCTATCATGACCCGAGTCTTTAGCTCAGGATAT 2658
Db      741 TrpValSerGlyValSerAenPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY      2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTCTTCTGATCATCGATGTTT 2718
Db      761 ArgTyrIleSerGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
QY      2719 GGTCTAGCATTTACCGAAGTATTGTGTAGATCTAAAGATTATGTAGTGTGCTTCCAAT 2778
Db      781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800
QY      2779 CATCATGCTTGCATAGATCCGTTTATCTATCTATCTACCCAAACAGCTTTATGTGATCCTAT 2838
Db      801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY      2839 TTGTTCCGAGATGCGTTTATCCGTCTAGCTACGGGTGTGGGATCAGCATATGAAACC 2898
Db      821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
QY      2899 TCATATACATTTCCAGAGAGACGATGTTCTGTTGGGATAATACTGCTCTGGCTGGAGAG 2958
Db      841 SerTyrThrPheAlaGluSerAspValArgTrpAspAenAenCysLeuValGlyGlu 860
QY      2959 ATTGGAGCGGGATTACCGATTGTGATTCTTCAATCTAAGCTCTATTGTAATGATGTCGT 3018
Db      861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
QY      3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGATCTTTTACAGAGGAGCGCAT 3078
Db      881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY      3079 CAAGCTCGGSCATTCAGAGCGGACATCTCTAAATCTATCAGTTCTGTTGGAGTGAAG 3138
Db      901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
QY      3139 TTTGATCGATGTTCTAGTACACATCCCTAATATAATAGCTTTTATGGCGGCTTATCTGT 3198
Db      921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
QY      3199 GATGCTTATCGACCATCTCTGCTAGTACAGACACGCTCTCTATCCCATCAAGACATGG 3258
Db      941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY      3259 ACAACAGATCCCTTTCATTTAGCAAGACATGGAGTTTGTGTGTAGGATCTATGTATGCT 3318
Db      961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY      3319 TCTTAAACAGTAATATAGAGTATATGGCCATCGGAAGATATAGTATCGAGATGCTTCT 3378
Db      981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY      3379 CGAGGCTATCGTTTGTAGTCAGGAAGTAGAGTCCGGTTC 3417
Db      1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

```

RESULT 17

US-10-701-844-15

```

; Sequence 15, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/512,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-15

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.4% Conservative: 17
Best Local Similarity: 96.7% Mismatches: 15
Query Match: 64.8% Indels: 1
DB: 4 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-701-844-15 (1-1013)

QY      382 ATGCAAACTCTTTCATAGTTCTTCTTCAATGATTCTAGCTTATTCTTCTCTCTCT 441
Db      1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCys 20
QY      442 TTAATATGGGGGGATATGCACAGAAATCATGGTTCCTCAAGAAATTTACGATGGGAG 501
Db      21 LeuAenGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY      502 AGTTAACTGTATCTTCCCTATCTACTGTATAGGAGATCCGAGTGGGACTACTGTTTTT 561
Db      41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY      562 TCTGCAGGAGAGTTAACTATAAAATCTTGCAATCTTATTCAGCTTTTGCCTTTAAGT 621
Db      61 SerAlaGlyGluLeuThrLeuLysAenLeuAenAenSerIleAlaLeuProLeuSer 80
QY      622 TGTTTTGGAACTTATTAGGAGTTTATCTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
Db      81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY      682 GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCCGCTGCTGATGACTG 741
Db      101 GluAenIleArgThrSerThrAenGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120
QY      742 TTTTACTATTGAGGTTTTTAAAGAAATTTATCTTTTCAATTTGCAATTCATTACTTGCCTA 801
Db      121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenProLeuLeuAlaVal 140
QY      802 CTCCTCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACACATCTACACCGTCT 861
Db      141 LeuProAlaAlaThrThrAenAenGlySerGlnThrProSerThrThrSerThrProSer 160
QY      862 AATGGTACTATTATTCTTAAACAGATCTTTTGTACTCAATATGAGAAGTTCTCATTC 921
Db      161 AenGlyThrIleTyrSerLysThrAspLeuLeuLeuAenAenGluLysPheSerPhe 180
QY      922 TATAGTAATTTACTCTCTGGAGATGGGGAGCTATAGATCTAAGAGCTTTAAACGGTTCAA 981
Db      181 TyrSerAenSerValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY      982 GGAATTAGCAAGCTTTGTGTCTTCCAAAGAAATATCTGCTCAAGCTGATGGGGAGCTGT 1041

```



Db 201 GlyIleSerLysLeuCyseValPheGlnGluAenThrAlaGlnAlaAaspGlyGlyAlaCys 220  
Qy 1042 CAAGTAGTACCAGTTCTCTGCTATGCTAACGAGGCTCCTATTGCTTTGTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheValAlaAen 240  
Qy 1102 GTTCGACGAGTAAGAGGGAGGATGCTGCTGTTTCAGGATGGGCACAGGAGTGCA 1161  
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAaspGlyGlnGlnGlyValSer 260  
Qy 1162 TCATCTACTCAACAGAAGATCCAGTAGTAAGTTTTTCCAGAAATACCTGCGGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAaspProValValSerPheSerArgAenThrAlaValGluPhe 280  
Qy 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCCTACGGAAACGTTGCTTCTG 1281  
Db 281 AspGlyAenValAlaArgValGlyGlyIleTySerTyGlyAenValAlaPheLeu 300  
Qy 1282 AATAATGAAAAACCTTGTTCTCAACAATGTTGCTTCTCTGTTTATCATGCTGCTAAG 1341  
Db 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyIleAlaAlaGlu 320  
Qy 1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATAATTACGAGATGGAGGAGCTATC 1401  
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAaspAenTyGlyAaspGlyAlaIle 340  
Qy 1402 TTCTGTAAAGATGGTGCACAA---GCAGGATCCAAATACTCTGGATCAGTTTCTTTCAT 1458  
Db 341 PheCysLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAasp 360  
Qy 1459 CGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGAAAGGGGAGCTATTTAT 1518  
Db 361 GlyGluGlyValAlaPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTy 380  
Qy 1519 GCCAAAAAGCTCTCGTTGCTAACTGTGCGCTGTACAAATTTTAAAGAAATACCTAAT 1578  
Db 381 AlalaLysLysLeuSerValAlaAenCysGlyProValGlnLeuLeuGlyAenIleAlaAen 400  
Qy 1579 GATGCTGAGCGCAATTATTATTAGAGAATCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638  
Db 401 AspGlyGlyAlaIleTyLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAaspTyGly 420  
Qy 1639 GATATTATTTTCAGTGGGAATCTTAAAGAACAGCAAGAGATGCTGCCGATGTTAT 1698  
Db 421 AspMetIlePheAaspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAaspValAen 440  
Qy 1699 GCGTAACCTGTCTCTCAAGCCATTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyValIleThrThrLeu 460  
Qy 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGCGCAACGAAAT 1818  
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAaspProIleUmetAlaAenGlyAen 480  
Qy 1819 AACGAGCGAGCGAGTCTTCCAAACTTCAAAAAATTACGATGGTGAAGGATACACAGG 1878  
Db 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAaspGlyGlyTyThrGly 500  
Qy 1879 GATATTGTTTTTGTAAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGA 1938  
Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyGlnAenValThrIleGlnGly 520  
Qy 1939 AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGG 1998  
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540  
Qy 1999 AGTCTGTATATGAAGCTGGAGTACATGGGATTTTGTAACTCCCAACACCACACAG 2058  
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAaspPheValThrProGlnProGlnGln 560  
Qy 2059 CTTCTGCGCGTAACTCAGTTGATCAGCTTTCCAATCTGCATTTGTCTCTTCTCTTTG 2118  
Db 561 ProProAlaAlaAenGlnSerIleThrLeuSerAenLeuHisLeuSerLeuSerLeu 580

Qy 2119 TTACCAACAATGAGTTCAGATTCCGAATCCTCTACCAATCTCTCCAGCGCAAGATTTCTCATCCT 2178  
Db 581 LeuAlaAenAenAlaValThrAenProThrAenProThrAenProAlaGlnAaspSerHisPro 600  
Qy 2179 GCAGTCAATTGCTAGCACAACTGCTGTTCTGTTCAATTAGTGGGCTTATCTTTTTCAG 2238  
Db 601 AlaValIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
Qy 2239 GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGGTCTTAATCAAAAAATCAAT 2298  
Db 621 AspLeuAaspThrAlaTyAspArgTyAspTrpLeuGlySerAenGlnLysIleAasp 640  
Qy 2299 GTCTCGAAAATTACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCGAGATTTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAenAlaProSerAaspLeuThrLeu 660  
Qy 2359 GGGATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGATCCTAAT 2418  
Db 661 GlyAenGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAaspProAen 680  
Qy 2419 ACAGCAATAATGCTCTTATCTCTGAAAGCTACATGAGCTAAACCTGGGTATAATCCT 2478  
Db 681 ThrAlaAenAenGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAenPro 700  
Qy 2479 GGGCTGAGCAGTAGCTTCTTTCGTTCCAAATAGTTTATGGGATCCATTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAaspIle 720  
Qy 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTTGTCGAGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAaspGlyArgSerTyCysArgGlyLeu 740  
Qy 2599 TGGGTTTCTGAGGATTTTCAATTTCTCTCATCATGACCGCGATGCTTTTAGTCAGGATAT 2658  
Db 741 TrpValSerGlyValSerAenPheTyRhiAaspArgAaspAlaLeuGlyGlnGlyTy 760  
Qy 2659 CGGTATATTAGTGGGGTTTCTTCTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718  
Db 761 ArgTyRhiLeSerGlyGlyTySerLeuGlyAlaAenSerTyRhiPheGlySerSerMetPhe 780  
Qy 2719 GGTCTAGCATTTACCAGAGTATTTGGTAGATCTAAAGATTATGATGTGCTGTTCCCAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAaspTyValValCysArgSerAen 800  
Qy 2779 CATCATGCTTGATAGGATCCGTTTATCTATCTACCCCAACAGCTTTATGTGATPCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyRleuSerThrLysGlnAlaLeuCysGlySerTy 820  
Qy 2839 TTGTTTCGAGATGCGTTTATCCGTCTAGCTACGGGTTTGGGAATCAGCATATGAAAACC 2898  
Db 821 ValPheGlyAaspAlaPheIleArgAlaSerTyRhiPheGlyAenGlnHisMetLysThr 840  
Qy 2899 TCATATACATTTGCGAGGAGGAGCGATGCTCGTTGGGATAATACTGCTGCTGCGAGAG 2958  
Db 841 SerTyThrPheAlaGluGluSerAaspValCysTrpAaspAenAenCysLeuValGlyGlu 860  
Qy 2959 ATTGAGCGGGAATTACCGATTGTGATTACTCCATCTAAGCTCTATTGTAATGAGTTGCGT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyRleuAenGluLeuArg 880  
Qy 3019 CTTTTCGCGAAGCTGAGTTTCTTATTCGCCGATCATGAAATCTTTTACAGAGAAAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyRhiAaspHisGluSerPheThrGluGluGlyAasp 900  
Qy 3079 CAAGCTCCGGCATTCAGAGCGGACATCTCCATAATCTATCATGTTCTGTTGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920  
Qy 3139 TTTGATCGATGTTCTAGTACACATCCTAATAAATAGCTTTTATCGCGCGCTTATATCTGT 3198  
Db 921 PheAaspArgCysSerSerThrHisProAenLysTyRhiSerPheMetGlyAlaTyIleCys 940



Db 481 AsnGlnProAlaGlnSerSerGluProLeuIleAsnAspGlyGluGlyThrGly 500  
Qy 1879 GATATTGTTTGGTAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGGA 1938  
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly 520  
Qy 1939 AGGATTGTTCTCGTAAAGGCAAAATATATCAGTGAATTCCTAAGTCAGACAGGTGG 1998  
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
Qy 1999 AGTCGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCACCAACCAACAG 2058  
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
Qy 2059 CCTCTGCCCTAAATCAGTTGATCAGCTTTCCAACTCGCATTTGTCTCTTCTTTG 2118  
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerSerLeu 580  
Qy 2119 TTAGCAACAATGCGATTACGAATCTCTTACCAATCTCTCCAGCGCAGATCTCATCT 2178  
Db 581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProThrAsnProAlaGlnAspSerHisPro 600  
Qy 2179 GCAGTCATTGGTAGCACAACTGCTGGTCTGTTCACAAATTAGTGGGCCCTATCTTTTCAG 2238  
Db 601 AlaValIleGlySerThrThrAlaGlyProValThrIleSerGlyProPhePheGlu 620  
Qy 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAAATCAAAAATCAAT 2298  
Db 621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsp 640  
Qy 2299 GTCCTGAAATTTACAGTTAGGACTAAGCCCGCCTAATGCCCCATCAGATTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
Qy 2359 GGGATGAGATGCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGTGGATCCTAAT 2418  
Db 661 GlyAsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
Qy 2419 ACAGCAATAATGGTCTTATCTACTCTGAAAGCTACATGACTTAAACCTGGGTATAATCCT 2478  
Db 681 ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTrpLysThrGlyTyAsnPro 700  
Qy 2479 GGGCTGAGAGTATGCTTTGTTGTTCCAAATAGTTTATGGGATCCATTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
Qy 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGGATGGCGCTCTTATTGTCGAGGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740  
Qy 2599 TGGGTTTCTGGAGTTTCGAATTTCTTATCATGACCGCGATGCTTTAGTCAGGATAT 2658  
Db 741 TrpValSerGlyValSerAsnPheSerTyHisAspArgAspAlaLeuGlyGlnGlyTy 760  
Qy 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTCTACTTTGGATCATCGATTT 2718  
Db 761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAsnSerTyPheGlySerMetPhe 780  
Qy 2719 GGTCTAGCATTTACCGAGTATTTGCTAGATCTAAAGATTATGATGTCGTTCCCAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAsn 800  
Qy 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTGGATCCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820  
Qy 2839 TTGTTCCGGAGATCGTTTATCCGTCTAGCTAGCTAGGTTTGGGAATCAGCATATGAAACC 2898  
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyGlyPheGlyAsnGlnHisMetLysThr 840  
Qy 2899 TCATATACATTTGCAGAGAGGACGATGTTCTGGGATAATTAATCTGCTGCGTCGAGAG 2958

Db 841 SerTyThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860  
Qy 2959 ATTGAGGGGATTACCGATTGCGATTCTCCATCTAAGCTCTATTGGAATGAGTTGGT 3018  
Db 861 IleGlyValGlyLeuProIleValThrThrProSerLysLeuTyLeuAsnGluLeuArg 880  
Qy 3019 CCTTTCTGTCAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy 3079 CAAGCTCGGCGATTCAAGAGCGGACATCTCCTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
Qy 3139 TTTGATGATGATCTTCTAGTACATCCTCTAAATAATATAGCTTTATGCGGCTTATCTGT 3198  
Db 921 PheAspArgCysSerSerThrHisProAsnLysTySerPheMetGlyAlaTyIleCys 940  
Qy 3199 GATGCTTATCCACCTCTCTGCTAGTACAGCAAGCTCCTATCCATCAAGAGACATGG 3258  
Db 941 AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGGAGTTGTGTTAGAGGATCTATGTATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyAla 980  
Qy 3319 TCTCTAACAAATATATAGAAATATATGCGCATGAGATATGAGTATGAGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyGlyHisGlyArgTyGluTyArgAspThrSer 1000  
Qy 3379 CGAGCTATGCTTTGAGTGCAGGAAGTAGATCCGGTTC 3417  
Db 1001 ArgGlyTyGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 19  
US-10-766-711-15  
; Sequence 15, Application US/10766711  
; Publication No. US20040137005A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-062-999  
; CURRENT APPLICATION NUMBER: US/10/766,711  
; CURRENT FILING DATE: 2004-01-27  
; PRIOR APPLICATION NUMBER: US/08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-766-711-15

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5106.50 Matches: 980  
Percent Similarity: 98.4% Conservative: 17  
Best Local Similarity: 96.7% Mismatches: 15  
Query Match: 64.8% Indels: 1  
Db: 4 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-766-711-15 (1-1013)

Qy 382 ATGCAACGCTCTTTCCATAAGTCTTTCTTTCAATGATTTAGCTTATTTCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTySerCysCysSer 20  
Qy 442 TTAATATGGGGGGGATATGCGACGAAATCATGTTCTCTCAAGGAATTTACATGGGGAG 501  
Db 21 LeuAsnGlyGlyTyArgTyAlaAlaGluIleMetValProGlnGlyIleTyAspGlyGlu 40

QY 502 AGCTTAACTGTATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGACTACTCTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGAGGAGAGTTAACTTAAATAAATCTTGCAAAATCTTATTGCGAGCTTTGCTTTAAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
QY 622 TGTTTTGGAACTTATTAGGAGTTTACTGTTTATAGGAGGAGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGAACATACGACTTCTCAAAATGGGCGAGCTCTAAGTAATAGCGCTCTGATGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
QY 742 TTTTACTATTGAGGTTTTAAAGAATTATCTTTTCCAATTGCAATTCATTCTGCGGTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140  
QY 802 CTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACACATCTACACGCTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrProSer 160  
QY 862 AATGGTACTATTATTCTAAACAGACTTTTGTGTACTCAATAATGAGAAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe 180  
QY 922 TATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAAGAGCTTAAACGGTCAA 981  
Db 181 TyrSerAsnSerValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
QY 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATAATCTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
QY 1042 CAAGTAGTCACCAAGTTTCTGCTATGGCTTAAGAGCTCTTATTCCTTTGAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240  
QY 1102 GTTGAGGAGTAGTAAGGGGAGGAGTGTCTGTCTGTTTCAAGATGGGAGGAGGTGTCA 1161  
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValAlaGlnAspGlyGlnGlnValSer 260  
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCGAAATATCTGCGGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
QY 1222 GATGGNACTAGCCGAGTAGGAGGAGGATTTACTCTTACCGGAGCTGTCTTCTCCTG 1281  
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
QY 1282 AATAATGGAAAACTTGTCTCAAAATGTTGCTTCTCCTCTTTTACATTTGCTGCTAAG 1341  
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320  
QY 1342 CAACCAACAAAGTGACAGGCTTCTAATAACAGTAATAATTACGAGATGAGGAGCTATC 1401  
Db 321 GlnProThrAsnGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
QY 1402 TTTCTGTAAGATGGTGGCAA---GCAGGATCCATAACTCTGGATCAGTTTCTTTGAT 1458  
Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
QY 1459 GGAGAGGAGTAGTTTCTTTTATGAGCAATGTAGCTCTGCGAAAGGGGAGCTATTAT 1518  
Db 361 GlyGluGlyValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380  
QY 1519 GCCAAAAGCTCTCGGTGCTTAATCTGCGCCCTGTACAATTTTTTAAGGAATATCGCTAAT 1578  
Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnLeuLeuGlyAsnIleAlaAsn 400  
QY 1579 GATGGTGGAGCGATTATTTTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638

Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
QY 1639 GATATTATTTTGGTGGAACTTTAAAGAACAGCCAAAGAGAAATGCTGCCGATGCTTAAT 1698  
Db 421 AspMetIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
QY 1699 GCGCTAACTGTGCTCTCAAAAGCCATTTTCGATGGATCGGGAGGAAATAACGACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
QY 1759 AGAGCTAAAGCGAGCATCAGATTCTCTTTAATGATCCCATCGAGATGCGAAACCGAAT 1818  
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
QY 1819 AACCAAGCAGCGCAGCTCTTCCAAACTTCTAAAAATTAACAGATGGTGAAGGATCACAGGG 1878  
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleLeuAsnAspGlyGluGlyTyrThrGly 500  
QY 1879 GATATTGTTTTTGTGTAAGCAGTACTTTGTACCAAAATGTTTACGATAGAGCAAGGA 1938  
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
QY 1939 AGGATTGTTCTTCTGTAAGGCAAAATTAATCAGTGAATTCCTAAGTCAGACAGGTGG 1998  
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
QY 1999 AGTCTGTATATGAGAGCTGGGAGTACATGGGATTTTGAATCTTCACTCCCAACACCAACAG 2058  
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
QY 2059 CCTCTGCGCGCTTAATCAGTTGATCAGCTTTCCTCAATCTGCATTTGTCTCTTCTTTG 2118  
Db 561 PropAlaAlaAsnGlnSerIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
QY 2119 TTAGCAAAACAATGCAAGTTACGAATCTCTACAAATCTCCAGCGCAAGATTCATCCT 2178  
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
QY 2179 GCAGTCATTTGTTAGCACACTCTGCTCTGTTCACAAATAGTGGGCTATCTTTTGTAG 2238  
Db 601 AlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
QY 2239 GATTTGGATGATACAGCTTATGATAGTATGATGCTAGTTCTAATCAAAAATCAAT 2298  
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspThrLeuGlySerAsnGlnLysIleAsp 640  
QY 2299 GTCCTGAAATTACAGTTAGGAGCTAAGCCCGCAGCTTAATGCCCATCAGATTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAsnAlaProSerAspLeuThrLeu 660  
QY 2359 GGGAAATGAGATGCTTAAGTATGGCTATCAGGAGCTGGAAGCTCGGTGGATCCCTAAT 2418  
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680  
QY 2419 ACAGCAAAATAGTCTCTTATCTCTGAAAGCTACATGAGTAAACTGGGTATAATCCCT 2478  
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrThrThrLysThrGlyTyrAsnPro 700  
QY 2479 GGGCCTGAGCGAGTAGCTTCTTTGTTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720  
QY 2539 CGATCTGGCGATTCAGCAATTCAGCAGTGTGATGGGCGCTTATTTGTCGAGGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
QY 2599 TGGGTTTCTGGAGTTCTCGAATTTCTTCTATCATGACGCGATGCTTTAGGTACAGGATAT 2658  
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
QY 2659 CGGTATATTAGTGGGGGTATTCTCTTAGGAGCAAACTCTTACTTTGGATCATCGATGTTT 2718

Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
Qy 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGACTAAAGATTATGTAGTGTGCTGCCAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
Qy 2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCCCAACAGCTTTATGGATCCCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
Qy 2839 TTGTTCCGAGATCGCTTTATCGTCTAGCTACGGGTTTGGAAATCAGCATATGAAACC 2898  
Db 821 ValPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
Qy 2899 TCATATACATTTGCAGAGAGCGATGTTCCGTTGGGATAATACTGCTGGCTGGAGAG 2958  
Db 841 SerTyrThrPheAlaGluSerAspValCysTyrAspAsnAsnCysLeuValGlyGlu 860  
Qy 2959 ATTGGAGCGGATTCACGATTTGATTAATCTACCTAAGCTTATTTGAATGAGTTGGT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
Qy 3019 CCTTTCCGCTGAGCTGAGTTTCTTATCCGATCATGAATCTTTTACAGAGGAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy 3079 CAAGCTCGGCGATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
Qy 3139 TTGTATCATGTTCTAGTACATCATCTTAATAATATAGCTTATGCGGCTTATATCTGT 3198  
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
Qy 3199 GATGCTTATCGACCATCTCTGGTACTGAGACACGCTCTATCCATCCATCAAGAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGGAGTGTGGTTPAGAGGATCTATGATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
Qy 3319 TCTCTAACAGTAATATAGAGTATATGCGCATGGAAGATATGATGATGATGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
Qy 3379 CGAGCTATGTTTCAGTGCAGAGTAGAGTCCGCTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

## RESULT 20

US-10-766-711-16  
; Sequence 16, Application US/10766711  
; Publication No. US20040137005A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-062-999  
; CURRENT APPLICATION NUMBER: US/10/766,711  
; CURRENT FILING DATE: 2004-01-27  
; PRIOR APPLICATION NUMBER: US/08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 1013  
; TYPE: PR1  
; ORGANISM: Chlamydia sp.  
US-10-766-711-16

Alignment Scores:

Pred. No.: 0 Length: 1013

Score: 5106.50 Matches: 980  
Percent Similarity: 98.2% Conservative: 15  
Best Local Similarity: 96.7% Mismatches: 17  
Query Match: 64.8% Indels: 1  
DB: 4 Gaps: 1  
US-10-701-844-1 (1-4435) x US-10-766-711-16 (1-1013)  
Qy 382 ATGCAACAGCTTTCCATAAGTCTTTCTTTCAATGATCTAGCTTATCTTGTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLeuPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20  
Qy 442 TTAATCGGGGGGATATGACAGCAATCATGCTCTCAAGCAATTTACGATGGGAG 501  
Db 21 LeuThrGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40  
Qy 502 ACGTTAACTGATCATTTCCCTATCTATCTATAGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTCGAGAGAGTTAAACATTAATAAATCTTGACAAATCTTATGCAAGCTTTGCTTTAAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
Qy 622 TGTTTTGGAACTTTATAGGAGGTTTACTGTTTATAGGAGGAGGACACTGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy 682 GAGAACATACGCACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
Qy 742 TTTACTATGAGGGTTTAAAGAATTATCTTTTCCAAATTCCAATTCATTACTTGCCTGA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
Qy 802 CTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACCAACATCTACACCGCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
Qy 862 AATGCTACTATTATTCTAAACAGACTCTTTTGTGTACTCAATATATGAGAAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnGlnLysPheSerPhe 180  
Qy 922 TATAGTAATATGATCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981  
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyThrIleAspAlaLysSerLeuThrValGln 200  
Qy 982 GGAATTAGCAAGCTTTGTGCTCTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
Qy 1042 CAAAGTAGTCACAGTTCTCTGCTATGCTTAACGAGGCTCTTATGCTTGTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240  
Qy 1102 GTTCAGAGTAAGAGGGGGAGGATTCCTGCTGTTCCAGGATGGCAGAGGAGTGTCA 1161  
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATATCTGCGGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
Qy 1222 GATGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTCTACGGAAACGTTCTTCTTCCTG 1281  
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
Qy 1282 AATAATGGAAAAACCTTGTCTCAACAATGTGTCTTCTCTGTTTACATTCGCTGTAAG 1341  
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaGlu 320  
Qy 1342 CAACCAACAGTGGACAGGCTTCTTAATACGAGTAATAATTACGAGATGGAGAGCTATC 1401

Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle 340  
 QY 1402 TTCTCTAAAGATGGTGGCGAA--GCAGGATCAATAACTCTGGATCAGTTTCCTTGTAT 1458  
 Db 341 PheCysLeuAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360  
 QY 1459 GGAGAGGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCGGAAAGGGGGAGCTATTTAT 1518  
 Db 361 GlyGluGlyValValPhePheSerAsnValAlaAlaGlySerGlyAlaIleTyr 380  
 QY 1519 GCCAAAGAGCTCTCGGTCTCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578  
 Db 381 AlaLeuLeuLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400  
 QY 1579 GATGTGGAGCGGATTTATTTAGAGAAATCTGGAGAGCTCAGTTTATCTCTGATATGGA 1638  
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
 QY 1639 GATATTATTTTCGATGGGAATCTTAAAGCAACAGCCAAAGAGAAATGCTCCGATGTTAAT 1698  
 Db 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLeuGluAsnAlaAlaAspValAsn 440  
 QY 1699 GCGTAACGTGTCTCACAGCCATTTCCGATGGGATCGGAGGGAGGAAATAACGACATTA 1758  
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrLeu 460  
 QY 1759 AGAGCTAAAGCGGCGATCAGATTCTCTTTAATGATCCCAATCAGATGGCAACGGAAT 1818  
 Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
 QY 1819 AACCAGCCAGCGAGTCTCCAAACTTCTAAATAATTAACGATCGTCAAGGATACACAGG 1878  
 Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500  
 QY 1879 GATATTGTTTCTTAATGAAGCAGTACTTCTTACCAAAATCTTACGATAGAGCAAGGA 1938  
 Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520  
 QY 1939 AGGATTTCTTCTCGTGAAGGCAAAATATATCAGTGAATCTCTAAGTCAGACAGGTGGG 1998  
 Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
 QY 1999 AGTCTGTATGGAAGCTGGGATGATGATGGATTTGTAACTCTCCACCAACACACACAG 2058  
 Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560  
 QY 2059 CCTCTGCGCTAATCAGTTGTATCAGCTTTCCAACTCTGCAATTTGTCTCTTCTTCTTTG 2118  
 Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580  
 QY 2119 TTAGCAACCAATGAGTTACGAATCTCTCTACCAATCCTCCAGCGCAAGATCTCATCCT 2178  
 Db 581 LeuAlaAsnAlaValThrAsnProThrAsnProThrAsnProAlaGlnAspSerHisPro 600  
 QY 2179 GCAGTCATGGTAGCACACTGTGTTCTGTACAAATGATGGCCCTATCTTTTGTAG 2238  
 Db 601 AlaValIleGlySerThrThrAlaGlyProValThrIleSerGlyProPhePheGlu 620  
 QY 2239 GATTTGGATCATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298  
 Db 621 AspLeuAspAspThrAlaIleTyrAspArgTyrAspIlePheLeuGlySerAsnGlnLysIleAsp 640  
 QY 2299 GTCTCGAAATTTACAGTTAGGAGCTTAAGCCCCCAGCTAATGCCCATCAGATTTGACTCTA 2358  
 Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660  
 QY 2359 GGGAAATCAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAGCTGGTGGGATCTCTAAT 2418  
 Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaIleTyrAspProAsn 680  
 QY 2419 ACAGCAATATGCTCTTATCTCTCAAGCTACATGAGTCAAGCTGAGTCAAGCTGAGTCA 2478

Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700  
 QY 2479 GGGCTCTGAGCAGTAGTCTCTTTGGTTCCTCAATAGTATGGGATCCATTTTAGATATA 2538  
 Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720  
 QY 2539 CGATCTGCGCATTCAGCAATTCAAGCAATGTGGATGGCGCTCTTTATTTGTGAGATT 2598  
 Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
 QY 2599 TGGGTTCTGGAGTTTCCGAATTTCTCTCATGACCGCGATGCTTTAGGTCAAGGATAT 2658  
 Db 741 TrpValSerGlyValSerAsnPheSerTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
 QY 2659 CGGTATATTAGTGGGGTATTCTCTTAGGAGCAAACTCTACTTTGGGATCATCGATGTT 2718  
 Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
 QY 2719 GGTCTAGCATTTTACCGAATTTTGGTAGATCTTAAAGATATGTAGTGTGTCTTCAAT 2778  
 Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
 QY 2779 CATCATGCTGTGATAGGATCCGTTTATCTCTATCCCAACAGCTTTATGTGATCCTAT 2838  
 Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
 QY 2839 TTGTTGGAGAGTCGTTTATCCGTCTAGCTACGGGTTTGGGAATCAGCATATGAAAACC 2898  
 Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
 QY 2899 TCATATACATTTTCAGAGAGAGCGATGTTGCTGGGATAATACTGTCTGCTGGTGAGAG 2958  
 Db 841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860  
 QY 2959 ATTGGAGCGGATTCACGATTTGATTTCTTCCATCTAAGCTCTTATTTGATGATGTCGT 3018  
 Db 861 IleGlyValGlyLeuProIleValThrThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
 QY 3019 CCTTTCTGCTCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGAT 3078  
 Db 881 PropheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
 QY 3079 CAAGCTCTGGGCAATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTGGAGTGAAG 3138  
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
 QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGCGGCTTATCTGT 3198  
 Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
 QY 3199 GATGCTTATCGCACCATCTCTGCTACTGAGACAACTCTCTATCCCATCAAGAGACATGG 3258  
 Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTyr 960  
 QY 3259 ACAACAGATGCTCTTCAATTTAGCAAGACATGAGTGTGTGTTAGAGATCTATGTATGCT 3318  
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
 QY 3319 TCTCTAACAGTAAATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGTCTTCT 3378  
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
 QY 3379 CGAGGCTATGTTTGGTGCAGGAGTGTAGAGTCCGGTTC 3417  
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 21

US-10-931-779-15  
 ; Sequence 15, Application US/10931779  
 ; Publication No. US20050048557A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; APPLICANT: Pace, John



**; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof**

: FILE REFERENCE: BP104

FILE REFERENCE: DE 104  
: CURRENT APPLICATION NUMBER: US/10/931.779

CURRENT AFFILIATION NUMBER: 03/1  
: CURRENT FILING DATE: 2004-09-01

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

;; PRIOR APPLICATION NUMBER: 09/  
PRIOR FILING DATE: 2008-04-03

PRIOR FILING DATE: 200-0-  
NUMBER OF SEQ. IN NOS. 43

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: P

; SEQ ID NO 15

; LENGTH: 10

; TYPE: PRT

; ORGANISM: ch.

**Alignment Scores:**

0	Length:	1013
5106.50	Matches:	980
98.4%	Conservative:	17
96.7%	Mismatches:	15
64.8%	Indels:	1
5	Gaps:	1
DB:		

US-10-701-844-1 (1-4435) X US-10-931-779-15 (1-1013)

Qy	382	ATGCAACGCTCTTCCATAAGTCTCTTCTTTCAATGATCTAGCTATTCTCTGCTGCTCT	441
Db	1	MetGlnThrSerPheHisLeuSerPheLeuSerMetIleLeuAlaTy-SerCysCysSer	20
Qy	442	TTAAATGGGGGGGATATGCACGAGAAATCATGGTTCTCTCAAGGAATTTACCATGGGAG	501
Db	21	LeuAsnGlyGlyTyThrAlaAlaGluIleMetValProGlnGlyIleTyThrAspGlyGlu	40
Qy	502	ACGTTAACTGTATCATTTCCCTATACTGTATAGGAGATCCGAGTGGGACTACTGTTTT	561
Db	41	ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe	60
Qy	562	TCTGCAGGAGAGTTAAACATTAAAAATCTTGA CAATCTATTGCAGCTTTGCTTTAAGT	621
Db	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
Qy	622	TGTTTTGGGAACCTATTAGGGAGTTTTACTGTTTTTAGGGAGAGGACACTCGTTGACTTTC	681
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
Qy	682	GAGAACATACGAGCTTCTACAAATCGGGCAGCTCTAAAGTAATAGCGCTCTCATGGAGTG	741
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
Qy	742	TTTACTATTGAGGGTTTTAAAGAAATTATCTCTTTTCCAAATTCGAAATTCATTACTTGC	801
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal	140
Qy	802	CTGCCTGCTGCAACGACTTAATAAGGGTAGCCAGACTCCGACGACAACTCTACACCGTCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrSerThrProSer	160
Qy	862	AATGCTACTATTATTCTAAACAGACTCTTTTGTTTACTCAATAATGAGAAGTTCTCATTC	921
Db	161	AsnGlyThrIleTyThrSerLysThrAspLeuLeuLeuAsnGlnLysPheSerPhe	180
Qy	922	TATAGTAATTTAGTCTCTCGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTC	981
Db	181	TyrSerAsnSerValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln	200
Qy	982	GGAATTAGCAAGCTTTGTCTTCCAGAAAATACTCTGCTCAAGCTGATGGGGAGCTTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
Qy	1042	CAAGTAGTCACAGTTTCTCTGCTATGGCTTAACGAGGCTCTTATTCCTTTGTAGGGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn	240
Qy	1102	GTTCGACAGAGTAAGAGGGGGAGGAGTTGCTGCTGTTTTCAGGATGGGCAGCAGGGAGTG	1161



Db 601 AlaValIleGlySerThrThraAlaGlySerValThrIleSerGlyProIlePheGlu 620  
Qy 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAATCAAAAATCAAT 2298  
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640  
Qy 2299 GTCTGAAATACAGTTAGGACTAAGCCCGAGCTTAATGCCCATCAGATTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProAlaAsnAlaProSerAspLeuThrLeu 660  
Qy 2359 GGGATGAGATGCCATGATGCTTATCAAGGAGCTGGAAGCTGGCGGATCCTAAT 2418  
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680  
Qy 2419 ACAGCAATATAGTCTTATCTGTAAGCTACATGAGCTAAACTGGGTATATCCT 2478  
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpLysThrGlyTyrAsnPro 700  
Qy 2479 GGGCTGAGCGAGTAGCTTCTTTGGTCCAAATAGTTATGGGATCCATTTTATAGATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720  
Qy 2539 CCAATCTGCGCAATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGAGATPA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740  
Qy 2599 TGGGTTCTGGAGTTTGGAAATTTCTTCTATCATGACCGGATGCTTTAGTCAAGGATAT 2658  
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760  
Qy 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAACTCTTCTTCTGATCATCGATGTTT 2718  
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780  
Qy 2719 GGTCTAGCAATTCACCAAGTATTGTTAGTATCTAAAGATTATGTAGTGTGTCCTCAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800  
Qy 2779 CATCATGTTGCATAGATCCGTTTATCTATCTACCAACAGCTTTATGTGATCTCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820  
Qy 2839 TTGTTGGAGATCGGTTTATCCGTGCTAGTACGGTTTGGGATCAGCATATGAAACC 2898  
Db 821 ValPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840  
Qy 2899 TCATATACATTTCCAGAGAGAGCGATGTTGCTGGGATAATTAATCTCTGCTGGCTGGAG 2958  
Db 841 SerTyrThrPheAlaGluSerAspValCysTrpAspAsnAsnCysLeuValGlyGlu 860  
Qy 2959 ATTGGAGCGGATTAACCGATTGTGATTCTCATCTAAGCTCTATTGTAATGATGCGGT 3018  
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880  
Qy 3019 CTTTCTGTCAGCACTGATTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT 3078  
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
Qy 3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTAAATCTATCACTCTGCTGCTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
Qy 3139 TTTGATCGATGTTCTAGTACACATCTCTTAATTAATATAGCTTTTATGGCGGTTTATCTGT 3198  
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
Qy 3199 GATGCTTATCGACCATCTCTGTTACTGACACAGCTCTCTATCCCATCAAGAGACATGG 3258  
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
Qy 3259 ACAACAGATCGCTTTCATTTAGCAAGACATGGAGTGTGTTAGAGATCTATGATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980

Qy 3319 TCTCTAACAGTAATATAGAAATATATGGCCATATATGAGTATGAGTATCGAGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
Qy 3379 CGAGCTATGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013  
RESULT 22  
US-10-931-779-16  
; Sequence 16, Application US/10931779  
; Publication No. US20050048557A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: BP104  
; CURRENT APPLICATION NUMBER: US/10/931,779  
; CURRENT FILING DATE: 2004-09-01  
; PRIOR APPLICATION NUMBER: 09/542,520  
; PRIOR FILING DATE: 200-04-03  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 16  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-931-779-16  
Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5106.50 Matches: 980  
Percent Similarity: 98.2% Conservative: 15  
Best Local Similarity: 96.7% Mismatches: 17  
Query Match: 64.8% Indels: 1  
DB: 5 Gaps: 1  
US-10-701-844-1 (1-4435) x US-10-931-779-16 (1-1013)  
Qy 382 ATGCAAAAGCTTCTCCATAGTCTTCTTCTTCAATGATTTCTAGCTTATTTCTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20  
Qy 442 TTAATATGGGGGGGATATGCAGCAAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501  
Db 21 LeuThrGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40  
Qy 502 ACCTTAACGTATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
Qy 562 TCTGCAGGAGATTAAATATAAATCTTGCAATCTTATTCAGCTTTGCTTTAAGT 621  
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80  
Qy 622 TGTTTTGGCACTTATAGGAGTTTACTGTGTTTAGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
Qy 682 GAGAACATACGGAATCTCAAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
Qy 742 TTTTACTATTGAGGGTTTTAAAGAAATTAATCTTTTCCAAATGTCATTTACTTACTTCCGTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
Qy 802 CTCCTGCTGCAACGATTAATAGGTAGCCAGCTCCGACGACCAATCTACACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrThrProSer 160  
Qy 862 AATGGTACTATTTTATTCTAAACACAGATCTTTTGTGTTACTCAATATAGAGAGTTCTCATTC 921



QY 3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTAATCTATCATGTTCTCTGTTGGAGTGAAG 3138  
Db GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920  
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAAATATAGCTTTATGGCGGCTTATATCTGT 3198  
Db PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940  
QY 3199 GATGCTTATCCACATCTCTGGTACTGAGACACAGCTTCTATCCATCAAGACATGG 3258  
Db AspaLatyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTTCTATTTAGCAAGACATCGAGTTGTTAGAGGATCTATGTATGCT 3318  
Db ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValAlaArgLysSerMetTyrAla 980  
QY 3319 TCTCTAACAGTAATATAGATATATGCGCATATGAGATATGAGATATGAGATGCTTCT 3378  
Db SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
QY 3379 CGAGCTATGTTGAGTGCAGGAAGTAGATCGGTTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

## RESULT 23

US-09-841-132-190  
; Sequence 190, Application US/09841132  
; Patent No. US20020061848A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhactia, Ajay  
; APPLICANT: Skeiky, Yahir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C8  
; CURRENT APPLICATION NUMBER: US/09/841.132  
; CURRENT FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 190  
; LENGTH: 1006  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-09-841-132-190

Alignment Scores:  
Pred. No.: 0 Length: 1006  
Score: 5090.00 Matches: 979  
Percent Similarity: 99.9% Conservative: 2  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 64.6% Indels: 0  
DB: 3 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-841-132-190 (1-1006)

QY 472 ATGGTTCTCAAGGAATTTACGATGGGAGAGCTTAACGTATCAATTCCTTACTGTGT 531  
Db 25 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 44  
QY 532 ATAGGATCCGAGTGGGACTACTGTTTTCTCGCAGGAGAGTTAATCAATTAATAATCTT 591  
Db 45 IleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64  
QY 592 GACAATCTTATTCAGCTTTCCTTAACTTGTGTTTGGGAACCTTATTAGGAGTTTACT 651  
Db 65 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84  
QY 652 GTTTATGGAGGAGGACATCGTTGACTTTTCGAGAACATACGAGCTTCTCAATGGGCA 711  
Db 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104  
QY 712 GCTCTAAGTAATAGCGCTCTGATGACTGTTTACTATTGAGGGTTTAAAGAAATATCC 771

Db 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124  
QY 772 TTTTCCAAATGCAATTCATTACTTGCCTGCTGCTCAACAGCATTAATAAGGTTAGC 831  
Db 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144  
QY 832 CAGACTCCGACGACAAACATCTACACCTCTAATAGTACTATTATTCTTAAACAGATCTT 891  
Db 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 164  
QY 892 TTGTTACTCAATATGAGAAGTTCTCATCTATAGTAAATTTAGTCTCTCGGAGATGGGGA 951  
Db 165 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 184  
QY 952 GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAAGAA 1011  
Db 185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 204  
QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACAGTTTCTCTGTATGCT 1071  
Db 205 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224  
QY 1072 AACGAGGCTCTATTGCTTGTAGCGAATGTTGCAGAGTAAAGGGGGAGGATTGCT 1131  
Db 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 244  
QY 1132 GCTGTTCCAGGATGGCAGCAGGAGTGTCACTCATCTACTTCAACAGAGATCCAGTAGTA 1191  
Db 245 AlaValGlnAspGlyGlnGlnGlyValSerSerThrThrSerThrGluAspProValVal 264  
QY 1192 AGTTTTCACAGAAATACTGCGGTAGAGTTTGTATGGAAACGTAGCCCGAGTAGGAGAGG 1251  
Db 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 284  
QY 1252 ATTTACTCTACGGGAACGTTGCTTCTGTAATATGAAAAACCTTGTCTTCAACAT 1311  
Db 285 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 304  
QY 1312 GTTGCTTCTCTGTTTACATTCGTCTAAGCAACCAACAGTGGACAGCTTCTAATACG 1371  
Db 305 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324  
QY 1372 AGTAATAATTACGAGATGGAGGAGCTATCTTCTGTAAGAATGGTGCAGCAGAGATCC 1431  
Db 325 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344  
QY 1432 AATAACTCTGATCAGTTTCTTTCATCGAGAGGAGTAGTTTCTTTTAGTAGCAATGTA 1491  
Db 345 AsnAsnSerGlySerValSerPheAspGlyGlyGluGlyValValPhePheSerSerAsnVal 364  
QY 1492 GCTGCTGGAAAGGGGAGCTATTTATGCCAAAAGCTCTCGTTGCTTAACCTGCGCCCT 1551  
Db 365 AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 384  
QY 1552 GTACAAATTTTAAAGGAATATCGCTAATGATGGTGGAGGATTTATTTAGGAGAATCTGA 1611  
Db 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGly 404  
QY 1612 GAGCTCAGTTTATCTGCTGATTCGAGATATTTATTTTCATGGGAATCTTAAAGAAACA 1671  
Db 405 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424  
QY 1672 GCCAAAGAGATGCTGCCGATGTTAATGCGGTAACTGTGTCTCTCAACAGCCATTCGATG 1731  
Db 425 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444  
QY 1732 CGATCGGAGGAGGAAATTAACGACATTAAGAGCTTAAAGCAGGAGGAGCATCAGATTCCTTAA 1791  
Db 445 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464  
QY 1792 GATCCCATCGAGATGGCAACCGGAATAAACCGACCGAGCGATCTTTCAAATCTTCAAAA 1851

465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484  
1852 ATTAACCATGTGAAGGATACACAGGCGATATTCTTTTGTCTAATGAAGCAGTACTTTG 1911  
485 IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504  
1912 TACCAAAATGTTACGATAGAGCAAGGAAGGATGTTCTTCTGTAAGGCAAAATATCA 1971  
505 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 524  
1972 GTGAATCTCTAAGTCACAGAGTGGGAGTCTGTATATGAGCTGGGAGTACATGGAT 2031  
525 ValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 544  
2032 TTTCTTAACCTCACAAACACACACAGCTCTCTCGCTAATCACTAGTTGATCACGCTTTC 2091  
545 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 564  
2092 AATCTGATTTGCTCTCTTTCTTTCTTTGTTGCAACAATGAGTTACGAATCTCTCCACC 2151  
565 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAlaValThrAsnProProThr 584  
2152 AATCTCTCAGCGCAAGATCTCATCTGCAGTCTTGTGTAGCACAACTGCTGTTCTGTT 2211  
585 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 604  
2212 ACAATTAGTGGCCTATCTTTTGTAGGATTTGGATGATACAGCTTATGATAGGTATGAT 2271  
605 ThrIleSerGlyProIlePhePheGluAspLeuAspThrAlaTyrAspArgTyrAsp 624  
2272 TGGCTAGGTTCTAAATCAAAAATCAATGCTCTGAAATTTACAGTTAGGAGCTTAAGCCCCA 2331  
625 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 644  
2332 GCTAATGCCCCATCGATTTGACTCTAGGGAATGAGATGCCCTAAGTATGGCTATCAAGGA 2391  
645 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 664  
2392 AGCTGGAAGCTTGGTGGGATCTTAATACAGCAAAATATGTCCTTACTCTGAAAGCT 2451  
665 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 684  
2452 ACATGGACTAAACCTGGGTATAATCTCGGCTGAGCAGTACTCTTTGGTTCCAAAT 2511  
685 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 704  
2512 AGTTATGGGATTCATTTAGATATACATCTGCGCATTCAGCAATTCAGCAATTCAGCAATG 2571  
705 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724  
2572 GATGGCGCTCTTATTTGTCGAGGATATGGGTTTCTGAGGTTTCGAAATTTCTCTCATCAT 2631  
725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPheTyrHis 744  
2632 CACCGCGATGCTTTAGTCAGGATATCGGTATATTAGTGGGGTTTATTCCTTAGAGACA 2691  
745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764  
2692 AACTCTACTTTGGATCATCGATGTTGTCAGCATTTACCAAGTATTTGGTATGATCT 2751  
765 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784  
2752 AAAGATTATGATGTGTCCTCAATCATCATGCTTGCATAGGATCCGTTTATCTATCT 2811  
785 LysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyrLeuSer 804  
2812 ACCCAACAGCTTTATGTGGATCCTATTGTTGCGAGATGCGTTTATCGCTAGCTAC 2871  
805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824  
2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCGAGGAGGAGGAGGATGTTGCT 2931  
825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 844

2932 TGGGATAATACTGTCTGCTGGCTGGAGAGATTGGAGCGGATTACCGATTGTGATTACTCCA 2991  
845 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864  
2992 TCTAAGCTCTATTGATGAGTTGCGTCTTTCGTCGAAGCTGAGTTTCTTATGCGCAT 3051  
865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884  
3052 CATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCAATTCAGAGCGGACATCTCCTA 3111  
885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904  
3112 AATCTATCAGTCTCTGTTGAGTGAAGTTTGTATGATGTTCTTAGTACACATCTTAATAA 3171  
905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924  
3172 TATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231  
925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944  
3232 ACGTCTCTATCCCATCAAGAGACATGCAACACAGATGCCCTTTTCATTTTAGCAAGCATGGA 3291  
945 ThrLeuLeuSerHisGlnGluThrTyrThrThrAspAlaPheHisLeuAlaArgHisGly 964  
3292 GTTCTGTTAGAGATCTATGATGTTCTCTTAAAGTATATAGAAATATATGAGTATATGGCAT 3351  
965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984  
3352 GGAAGATATGATGATCGATGCTTCTCGAGGCTATGTTGAGTCAGGAGGATAGAGTC 3411  
985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerIleVal 1004  
3412 CGGTTCT 3417  
1005 ArgPhe 1006  
RESULT 24  
US-10-872-155-190  
; Sequence 190, Application US/10872155  
; Publication No. US20040234536A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C10  
; CURRENT APPLICATION NUMBER: US/10/872,155  
; PRIOR FILING DATE: 2004-06-18  
; PRIOR APPLICATION NUMBER: 09/620,412  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 09/598,419  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/556,877  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: 09/454,684  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: 09/426,571  
; PRIOR FILING DATE: 1999-10-22  
; PRIOR APPLICATION NUMBER: 09/410,568  
; PRIOR FILING DATE: 1999-10-01  
; PRIOR APPLICATION NUMBER: 09/288,594  
; PRIOR FILING DATE: 1999-04-08  
; PRIOR APPLICATION NUMBER: 09/208,277  
; PRIOR FILING DATE: 1998-12-08  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 190  
; LENGTH: 1006  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-10-872-155-190

## Alignment Scores:

Pred. No.: 0  
 Score: 5090.00  
 Percent Similarity: 99.9%  
 Best Local Similarity: 99.7%  
 Query Match: 64.6%  
 DB: 5

Length: 1006  
 Matches: 979  
 Conservative: 2  
 Mismatches: 1  
 Indels: 0  
 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-872-155-190 (1-1006)

QY 472 ATGTTCTCTCAAGGAATTTACGATGGGAGAGCTTAACATGTATCATTTCCCTTACTGTT 531  
 DB 25 MetileProGlnGlyIleTyAspGlyGluThrLeuThrValSerPheProTyThrVal 44  
 QY 532 ATAGAGATCCGAGTGGGACTACTGTTTTCGCGAGGAGGTAAACATTAATAATCTT 591  
 DB 45 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64  
 QY 592 GACAAATCTATTCGACGCTTTCCTTAACTTCTTTGGGAATTTATAGGAGTTTACT 651  
 DB 65 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84  
 QY 652 GTTTTAGGAGAGGACACCTCGTTGACTTTTCGAGAACATACGGACTTCTACAAATGGGCA 711  
 DB 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104  
 QY 712 GCTCTAAGTAATAGCGCTCTCATGAGCTGTTTATCTATTGAGGGTTTAAAGATATCC 771  
 DB 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124  
 QY 772 TTTTCCAATTGCAATTCATTCTGCGTACTGCTGCTGCAACGACTAATAAGGTAGC 831  
 DB 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144  
 QY 832 CAGACTCCGACGACACATCTACACCGTCTAATGGTACTATTATTCTAAACACAGATCTT 891  
 DB 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTySerLysThrAspLeu 164  
 QY 892 TTGTACTCAATAATAGAGAGTTCTCATCTATAGTAATTTAGTCTCTGAGATGGGGA 951  
 DB 165 LeuLeuLeuAsnAsnGlyLysPheSerPheTySerAsnLeuValSerGlyAspGlyGly 184  
 QY 952 GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGTTGCTCTTCCAGAA 1011  
 DB 185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 204  
 QY 1012 AATACTGCTCAAGCTCATGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCT 1071  
 DB 205 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224  
 QY 1072 AACGAGGCTCTATTGCTTTGTAGCGAATGTTGCGAGAGTAAGAGGGGAGGATGCT 1131  
 DB 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 244  
 QY 1132 GCTGTTCCAGATGGGCGAGGAGTGTATCATCTACTTCAACAGAGATCCAGTAGTA 1191  
 DB 245 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrSerThrGluAspProValVal 264  
 QY 1192 AGTTTTTCCAGAAATACTCGGTAGTAGTTGTGAGGAACCTAGCCCGAGTAGGAGAGGG 1251  
 DB 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly 284  
 QY 1252 ATTTACTCTTACGGGAGCTTGTCTTCTCGAATAATGGAATAACCTTTGTTTCTCAACAAT 1311  
 DB 285 IleTySerTyGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 304  
 QY 1312 GTTGCTTCTCTGTTTACATTCGTCTAAGCAACCAACAGAGTGAAGGCTTCTAATACG 1371  
 DB 305 ValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324  
 QY 1372 AGTAATAATTCGAGATGAGGAGCTATCTTCTGTAAAGATGCTGCGCAAGCAGGATCC 1431

DB 325 SerAsnAsnTyGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344  
 QY 1432 AATAACTCTGATCAGTTTCTCTTATGCGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
 DB 345 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 364  
 QY 1492 CTTCTGCGAAGGGGAGCTATTATATGCCAAAAAGCTCTCGTTGCTTAACCTGCGCCT 1551  
 DB 365 AlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCysGlyPro 384  
 QY 1552 GTACAAATTTTAAAGAAATATCGTAATCGTGGAGGAGTATTATTTAGGAGAATCTGA 1611  
 DB 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyLeuGlyGluSerGly 404  
 QY 1612 GAGCTCAGTTTCTGCTGATTATGAGATATTTATTTTCGATGGGAATCTTAAAAAACA 1671  
 DB 405 GluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424  
 QY 1672 GCCAAGAGAATGCTGCGGATGTTAATCGGTAACATGTGTCTCACAGCCCATTTTCGATG 1731  
 DB 425 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444  
 QY 1732 CGATCGGAGGGAATAATACGACATTAGAGCTTAAGCAGGCGCATCAGATTCTCTTAAT 1791  
 DB 445 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464  
 QY 1792 GATCCCATCGAGATCGCAAAATTAACAGCCAGCGCAGCTTCCAAACCTCTTAAAA 1851  
 DB 465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484  
 QY 1852 ATTAACGATGTTGAAGGATACACAGGGATATGTTTTTGTCTTAATGGAAGAGTACTTGG 1911  
 DB 485 IleAsnAspGlyGluGlyTyThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504  
 QY 1912 TACCAAAATGTTACATAGACAGCAAGAGGATGTTCTTCGTGAAAGGCAAAATATCA 1971  
 DB 505 TyGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 524  
 QY 1972 GTGAATCTCTAAGTCAGACAGCTGCGAGTCTGTATATGGAAGCTGGGAGTAGTACATGGAT 2031  
 DB 525 ValAsnSerLeuSerGlnThrGlySerLeuTyMetGluAlaGlySerThrLeuAsp 544  
 QY 2032 TTTGTAACCTCCAAACCAACCAACAGCTCTCGCGCTTAATCAGTTGATCAGCTTTCC 2091  
 DB 545 PheValThrProGlnProGlnProAlaAlaAsnGlnLeuIleThrLeuSer 564  
 QY 2092 AATCTGATTTCTCTTCTTCTTCTTTAGCAAAATGCAAGTACAGTACGATCCTCTACC 2151  
 DB 565 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProThr 584  
 QY 2152 AATCTCCAGCGCAAGTTCTCATCTGCGAGTCATTGTTAGTAGCACAACTGCTGTTCTGTT 2211  
 DB 585 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 604  
 QY 2212 ACATATAGTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2271  
 DB 605 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyAspArgTyAsp 624  
 QY 2272 TGCTAGGTTCTTAATCAAAAAATCAATGTCTCGTGAATAATACAGTTAGGAGTACAGCCCA 2331  
 DB 625 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 644  
 QY 2332 GCTAATGCCCCATCAGATTTGCTCTAGGGAATGAGATGCTAAGTATGCTATCAAGGA 2391  
 DB 645 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyGlyTyGlnGly 664  
 QY 2392 AGCTGGAAGCTTGGGCTGCGGATCCATAACAGCAAAATAATGGTCTTATCTCTGAAGCT 2451  
 DB 665 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyThrLeuLysAla 684  
 QY 2452 ACATGGAATAAACTGGGTATTAATCTCGGCTGAGGAGTAGCTCTTCTTGGTTCGAAT 2511  
 DB 685 ThrTrpThrLysThrGlyTyAsnProGlyProGluArgValAlaSerLeuValProAsn 704

```
QY 2512 AGTTATCGGATCCATTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTG 2571
Db 705 SerLeuTyrGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
QY 2572 GATGGGCGCTTATTTGTCGAGGATATAGGTTCTCGAGTTTCGAAATTTCTTCATCAT 2631
Db 725 AspGlyArgSerTyrCysArgGlyLeuTyrPheValSerGlyValSerAsnPhePheTyrHis 744
QY 2632 GACCGCATGCTTTAGTTCAGGATATCGGTATATTAGTGGGGTTATTCTTAGGAGCA 2691
Db 745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764
QY 2692 AACTCCTACTTTGGATCATCGATTTGGTCTAGCATTTACGAAGTATTTGGTAGATCT 2751
Db 765 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
QY 2752 AAAGATTATGATGTGTGCTTCCATCATCATGCTTCATAGGATCGTTTATCTATCT 2811
Db 785 LysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyrLeuSer 804
QY 2812 ACCCAACAGCTTTATGTGGATCCTATTGTTGGGAGATGCGTTTATCCGTCTAGCTAC 2871
Db 805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824
QY 2872 GGGTTGGGAATCAGCATATGAAACCTCATATACATATTGTCGAGGAGCGATGTTGCT 2931
Db 825 GlyPheGlyAsnGlnHisMetIysThrSerTyrThrPheAlaGluGluSerAspValArg 844
QY 2932 TGGGATAATAACTGTCGTCGAGGATGAGGAGGATACCGGATTAACCATGTGATTAACCTCA 2991
Db 845 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864
QY 2992 TCTAAGCTCTATTGAATGAGTTGGCTCTTTCGTCGCAAGCTGAGTTTCTTATCCGAT 3051
Db 865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884
QY 3052 CATGAATCTTTTACAGAGAGCGCATCAAGCTCGGGCATTCAGAGCGGACATCTCCTA 3111
Db 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904
QY 3112 AATCTATCAGTTCCTGTTGGAGTGAAGTTTGATCGATGCTTCAGTACACATCCTAATAA 3171
Db 905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924
QY 3172 TATAGCTTTATCGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTACTGAGACA 3231
Db 925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944
QY 3232 ACGCTCTATCCCATCAGAGACATGGACAACAGATGCGCTTTCATTTAGCAAGATGGA 3291
Db 945 ThrLeuLeuSerHisGlnGluThrTrpThrAspAlaPheHisLeuAlaArgHisGly 964
QY 3292 GTTGTGGTTAGAGGATCTATGATGCTTCTCTAAAGTAAATATAGAGTATATGCCAT 3351
Db 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984
QY 3352 GAAAGATGATGATTCGAGATGCTTCTCGAGGCTATGTTTGGTGCAGGAAGTAGAGTC 3411
Db 985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004
QY 3412 CGGTTC 3417
Db 1005 ArgPhe 1006
```

## RESULT 25

US-09-841-132-176

; Sequence 176, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841.132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: PastSeq for Windows Version 3.0/4.0

; SEQ ID NO 176

; LENGTH: 982

; TYPE: PRT

; ORGANISM: Chlamydia

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(982)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-841-132-176

Alignment Scores:

Pred. No.: 0 Length: 982

Score: 5084.00 Matches: 978

Percent Similarity: 99.8% Conservative: 2

Best Local Similarity: 99.6% Mismatches: 2

Query Match: 64.5% Indels: 0

DB: 3 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-841-132-176 (1-982)

```
QY 472 ATGGTTCCTCAAGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATATCTGT 531
Db 1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20
QY 532 ATAGAGATCCGAGTGGGACTACTGTTTTTCTCGAGGAGATTAACTTAAAAATCTT 591
Db 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluThrLeuLysAsnLeu 40
QY 592 GACAAATCTATTTCAGCTTTCCTTAAAGTTGTTTTGGAACTTATTAGGGAGTTTACT 651
Db 41 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60
QY 652 GTTTTAGGGAGAGACACTCGTTGACTTTCGAGAACATACGACTTCTACAAATGGGCA 711
Db 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80
QY 712 GCTCTAAGTAAATAGCGCTGCTGATGGACTGTTTACTATTAGGGTTTAAAGAAATTATCC 771
Db 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGlyGlyPheLysGluLeuSer 100
QY 772 TTTTCCAATTGCAATTCAATTTCAATTTCTGCTGCTGCTGCAACGACTTAATAAGGGTACG 831
Db 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaIleThrAsnLysGlySer 120
QY 832 CAGACTCCGACGACAACTCTACACCGCTCTAATGGTACTATTATTCTTAAACACAGATCTT 891
Db 121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 140
QY 892 TTGTTACTCAATAATAGAGAGTTCTCATCTCTATAGTAATTTAGTCTCTCGAGATCGGGA 951
Db 141 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160
QY 952 GCTATAGATGCTTAAGAGCTTAAACGGTTCAAGGAATTTAGCAAGCTTTGTGTTCTTCAAGA 1011
Db 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuLysValPheGlnGlu 180
QY 1012 AATACTGCTCAAGCTGATGGGAGCTGTGTCAAGTAGTCACAGATTTCTCTGCTATGGCT 1071
Db 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200
QY 1072 AACGAGGCTCCTATTGCTTTGATCGGAATGTTGAGGAGTAAAGGGGAGGAGTGTCT 1131
Db 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 220
QY 1132 GCTGTTTCAGGATGGGACGAGGTGTCATCTACTTCAACAGAGATCCAGTAGTA 1191
```



Db 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrSerThrGluAspProValVal 240  
QY 1192 AGTTTTTCCAGAAATACCTCGGTAGAGTTTGATGGAACTAGCCGAGTAGAGAGGG 1251  
Db 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 260  
QY 1252 ATTTACTCTACGGAGCGTCTCTTCCCTGAATATGGAATAACCTTTCTCAACAAT 1311  
Db 261 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyTyrThrLeuPheLeuAsn 280  
QY 1312 GTTGTCTTCTCTGTGTACATTTGCTGAAGCAACCAAGTGGACAGGCTTCTAATACG 1371  
Db 281 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300  
QY 1372 AGTAATAATTACGGAGAGGAGCTATCTCTGTAAAGATGCTGCGGACAGGATCC 1431  
Db 301 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320  
QY 1432 AATAACTCTGGATCAGTTCTCTTGTATGGAGGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
Db 321 AsnAsnSerGlySerValSerPheAspGlyGlyValValPhePheSerSerAsnVal 340  
QY 1492 GCTGCTGGGAAAGGGGAGCTATTTATGCCAAAAGCTCTCGGTGCTAACTGTGCCCC 1551  
Db 341 AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 360  
QY 1552 GTACAAATTTTAAAGGAATATCGTAATGATGGTGGAGCGGATTTATTAGGAGAACTGGA 1611  
Db 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlySerGly 380  
QY 1612 GAGCTCAGTTTATCTCTGATTTAGGATATTTATTTCCATCGGATCTTAAAGAACCA 1671  
Db 381 GluLeuSerLeuSerAlaAspTyrGlyAspIlePheAspGlyAsnLeuLysAsnGthr 400  
QY 1672 GCCAAAGAGAACTGCTGCCGATTTAATGGCGTAATCTGTGCTCCTCACAAAGCCATTTGATG 1731  
Db 401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420  
QY 1732 GGATCGGGAGGGAATAATACACATTAAGAGCTAAAGCAGCGGCATCAGATCTCTTTAAT 1791  
Db 421 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyIleGlnIleLeuPheAsn 440  
QY 1792 GATCCCATCAGATGGCAACCGGAAATAACAGCCAGCGCAGCTTCCAACTCTTAAAA 1851  
Db 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerLysLeuLys 460  
QY 1852 ATTAACGATGCTGAAGGATACACAGGGATTTCTTTTGTATGTAAGCAGTACTTTG 1911  
Db 461 IleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480  
QY 1912 TACCAAAATGTTACGATAGACGAAGGAGTGTCTTCTCGTGAAGGCAAAATTAACA 1971  
Db 481 TyrGlnAsnValThrIleGluGlnArgIleValLeuArgGluLysAlaLysLeuSer 500  
QY 1972 GTGAATTTCTTAAGTACAGAGTGGAGTGTATATGAGTGGAGTGGAGTACATGGAT 2031  
Db 501 ValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 520  
QY 2032 TTTGTAACTCCACACACCAACAGCCTCTCGCGCTAATCAGTTGTATGATCAGCCTTCC 2091  
Db 521 PheValThrProGlnProProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 540  
QY 2092 AATCTGCAATTTGTCTCTTCTTCTTTGTAGCAACAATGAGTTACGAATPCTCTTACC 2151  
Db 541 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProThr 560  
QY 2152 AATCTCCAGCGCAGATTTCTATCTCTGAGTGTGGTAGCAACATGCTGGTCTGTT 2211  
Db 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580  
QY 2212 ACAATTAGTGGCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGAT 2271  
Db 581 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 600

QY 2272 TGGCTAGGTTCTTAATCAAAAAATCAATGTCCTGAAATTTACAGTTAGGAGCTAAGCCCCA 2331  
Db 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620  
QY 2332 GCTAATGCCCATCAGATTTGACTTAGGAATAGAGCCCTAAGTATGGCTATCAGCA 2391  
Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 640  
QY 2392 AGCTGGAAGCTGCGTGGGATCCTAATACAGCAATAATGGTCTTATATCTCTGAAAGCT 2451  
Db 641 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660  
QY 2452 ACATCGAGCTAAAACTGGGTATTAATCCTGGGCTCGAGCGAGTAGCTTCTTTGGTCCAAAT 2511  
Db 661 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680  
QY 2512 AGTTATGGGATCCATTTTAGATATACGATCTGGCATTCAGCAATTCAGCAAGTGG 2571  
Db 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700  
QY 2572 GATGGCGCTCTTATTTGTCGAGGATTTATGGGTTCTCGAGTTTCGAATTTCTTCTATCAT 2631  
Db 701 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720  
QY 2632 GACCGCATGCTTTAGTTCAGGATATCGGTATATTAGTGGGGTTATTCTTTAGGAGCA 2691  
Db 721 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740  
QY 2692 AACTCTCTCTTTGGATCATCGATTTGCTAGCATTTTACCGAATTTTGGTAGATCT 2751  
Db 741 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760  
QY 2752 AAGATTATGATGTGTCTGTTCCATCATCATCTGCTTCGATAGTAGGTCGTTTATCTATCT 2811  
Db 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780  
QY 2812 ACCCAACAGCTTTATGTGATCCTATTGTTTCGAGATGCGTTTATTCGCTGCTAGCTAC 2871  
Db 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800  
QY 2872 GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTTCAGAGAGCGGATGTCGT 2931  
Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820  
QY 2932 TGGGATTAATACCTGCTGCTGAGAGATGGAGCGGATTCGATTCGATGATTACTCCA 2991  
Db 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840  
QY 2992 TCTAAGCTCTATTGATGATGTTGCTGCTTCTGTCGCAAGCTGAGTTTCTTATGCCGAT 3051  
Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860  
QY 3052 CATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCTTA 3111  
Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880  
QY 3112 AATCTATCAGTTCTGTTGGAGTCAAGTTTGATCGATGTTCTAGTACACATCCTAATAAA 3171  
Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900  
QY 3172 TATAGCTTTATGGCGCTTATATCTGTGATGCTTTATCGCACCATCTCTGGTACTGAGACA 3231  
Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920  
QY 3232 ACGCTCTATCTCCATCAAGAGACATGACCAACAGATCGCTTCTTATTTAGCAAGACATGGA 3291  
Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940  
QY 3292 GTTGTGTTAGAGGATCTATGATGCTTCTCTAAAGATATATAGAGTATATATGGCAT 3351  
Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960



QY 3352 CGAGATATGATCTCGAGATCTCTCGAGCTATGCTTGGAGTGCAGGAGTAGATC 3411  
Db |||||  
QY 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerIysVal 980  
3412 CGGTTCT 3417  
Db |||||

Db 981 \*\*\*Phe 982

## RESULT 26

US-10-872-155-176  
; Sequence 176, Application US/10872155  
; Publication No. US20040234536A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C10  
; CURRENT APPLICATION NUMBER: US/10/872,155  
; CURRENT FILING DATE: 2004-06-18  
; PRIORITY APPLICATION NUMBER: 09/620,412  
; PRIORITY FILING DATE: 2000-07-20  
; PRIORITY APPLICATION NUMBER: 09/598,419  
; PRIORITY FILING DATE: 2000-06-20  
; PRIORITY APPLICATION NUMBER: 09/556,877  
; PRIORITY FILING DATE: 2000-04-19  
; PRIORITY APPLICATION NUMBER: 09/454,684  
; PRIORITY FILING DATE: 1999-12-03  
; PRIORITY APPLICATION NUMBER: 09/426,571  
; PRIORITY FILING DATE: 1999-10-22  
; PRIORITY APPLICATION NUMBER: 09/410,568  
; PRIORITY FILING DATE: 1999-10-01  
; PRIORITY APPLICATION NUMBER: 09/288,594  
; PRIORITY FILING DATE: 1999-04-08  
; PRIORITY APPLICATION NUMBER: 09/208,277  
; PRIORITY FILING DATE: 1998-12-08  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 176  
; LENGTH: 982  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(982)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-872-155-176

Alignment Scores:  
Pred. No.: 0 Length: 982  
Score: 5084.00 Matches: 978  
Percent Similarity: 99.8% Conservative: 2  
Best Local Similarity: 99.6% Mismatches: 2  
Query Match: 64.5% Indels: 0  
DB: 5 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-872-155-176 (1-982)

QY 472 ATGGTCTCCAGGAAATTCAGATCGGAGAGCTTAACCTGATCATTTCCCTACTGTT 531  
Db |||||  
1 MetIleProGlnGlyIleTyrAspGlyGluThrValPhePheProTyrThrVal 20  
QY 532 ATAGAGATCCGAGTGGGACTACTGTTTTTCTCGAGGAGAGTTAAACATTAAAAATCTT 591  
Db |||||  
21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuIysAsnLeu 40  
QY 592 GACAAATCTATTCGAGCTTGTGCTTAACTGTTTGGAACTTATAGGAGTTTACT 651  
Db |||||  
41 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuGlySerPheThr 60  
QY 652 GTTTTAGGAGGAGACACTCGTTGACTTTCGAGACATACGACTTCTACAAATGGGCA 711  
Db |||||

Db 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80  
QY 712 GCTCTAAGTAATACGGCTGCTGATGAGTCTTTACTATTAGAGGTTTTTAAAGAATTATCC 771  
Db |||||  
81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 100  
772 TTTTCCAATTGCAATTCATTACTTCCGCTGCTGCTGCAACGACTTAATAAGGATAGC 831  
Db |||||  
101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrAsnLysGlySer 120  
QY 832 CAGACTCCGACGACAAATCTACACCGTCTAATGGTACTATTATTCTTAAACAGATCTT 891  
Db |||||  
121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 140  
QY 892 TTGTTACTCAATAATAGAGAGTCTCTCATCTTATAGTAAATTTAGTCTCTCGAGATGGGGA 951  
Db |||||  
141 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160  
QY 952 GCTATAGTGTCTAAGAGCTTAAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCCAAGAA 1011  
Db |||||  
161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180  
QY 1012 AATCTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTCTTCTCTGCTATGGCT 1071  
Db |||||  
181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200  
QY 1072 AACGAGGCTCTATTGCTTTGTAGCGAATTTTGCAGAGTAAAGAGGGGAGGATGCT 1131  
Db |||||  
201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 220  
QY 1132 GCTGTTCCAGATGGGACGAGGAGTGTCTATCTACTTCAACAGAGATCCAGTAGTA 1191  
Db |||||  
221 AlaValGlnAspGlyGlnGlnGlyValSerSerThrThrGluAspProValVal 240  
QY 1192 AGTTTTTCCAGAAATCTCGCGTAGAGTTTCAATGGGAACGTAGGCCGAGTAGGAGAGGG 1251  
Db |||||  
241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly 260  
QY 1252 ATTTACTCTACGGGAACGTTGCTTCTGTAATTAATGGAATAAACCCTTTCTCAACAAAT 1311  
Db |||||  
261 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 280  
QY 1312 GTTGCTTCTCTGTTTACTGCTTAAGCAACCAACAGCTGACAGGCTTCTTAATACG 1371  
Db |||||  
281 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGluAlaSerAsnThr 300  
QY 1372 AGTAATAATTACGGAGATGGAGGAGCTATCTTCTGTAAGAATGGTGCACAGCAGATCC 1431  
Db |||||  
301 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320  
QY 1432 AATAACTCTGGATCAGTTTCTTTGATGAGAGGAGTAGTTTCTTTTAGTAGCAATGTA 1491  
Db |||||  
321 AsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSerAsnVal 340  
QY 1492 GCTGCTGGAAAGGGGAGCTATTATGCCAAAAGCTCTCGGTGCTAACTGTGCCCT 1551  
Db |||||  
341 AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 360  
QY 1552 GTACAAATTTTAAAGGAATATCGTAAATGCTGAGGAGCATTTATTTAGGAGAACTGGA 1611  
Db |||||  
361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGly 380  
QY 1612 GAGCTCAGTTTATCTGCTGATTTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACA 1671  
Db |||||  
381 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400  
QY 1672 GCCAAGAAGATGCTGCCGATTTAATGCGCTAACTGTCTCTCAAGCCATTTCCGATG 1731  
Db |||||  
401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420  
QY 1732 GGATCGGAGGGGAAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATCTCTTTAAT 1791  
Db |||||  
421 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440



Db 1 AlaGluIleMetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPhePro 20  
Qy 523 TATACGTGTATAGGAGATCCGAGTGGGACTACTGTTTTTCTGCGAGGAGAGTTAAACATTA 582  
Db 21 TyrThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeu 40  
Qy 583 AAAATCTTGACAACTTATTGCAGCTTGGCTTTAACTTGGTGGGAACTTATTAGG 642  
Db 41 LysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGly 60  
Qy 643 AGTTTTACTGTTTTAGGAGGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACA 702  
Db 61 SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr 80  
Qy 703 AATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGAGGTTTTAAA 762  
Db 81 AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys 100  
Qy 763 GNAATTATCTTTTCCAAATTCGAATTCATTAATCTACTTGGCGTACTGCCCTGCGCAACGACTAAT 822  
Db 101 GluLeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsn 120  
Qy 823 AAGGGTAGCCAGACTCCGAGGACAACTCTACACCGTCTAAATGGTACTATTATTCTAAA 882  
Db 121 AsnGlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTyrSerLys 140  
Qy 883 ACAGATCTTTTGTACTCAATATAGAAAGTTCTCATTTCTATPAGTAATTTAGTCTCTGGA 942  
Db 141 ThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGly 160  
Qy 943 GATGGGCGAGCTATAGATCTTAAGACTTAACGGTTCAAGGAATTAGCAAGCTTTGTGTC 1002  
Db 161 AspGlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysVal 180  
Qy 1003 TTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACACAGTTTCTCT 1062  
Db 181 PheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSer 200  
Qy 1063 GCTATGGCTAACAGGCTCTATTGCTTTGTAGCGAATGTTGCGAGGAGTAAGAGGGGA 1122  
Db 201 AlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsnValAlaGlyValArgGlyGly 220  
Qy 1123 GGGATGCTGCTGTTCCAGATGGCGAGGAGGTGTCATCTACTTCAACAGAGAT 1182  
Db 221 GlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrSerThrGluAsp 240  
Qy 1183 CCAGTAGTAAGTTTTCCAGAAATACTCGGTAGAGTTTGTAGGAACTAGCCGAGTA 1242  
Db 241 ProValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgVal 260  
Qy 1243 GGAGGAGGATTTACTCTACGGGAACGTTGCTTTCTCTGAAATATGGAATAACCTTGTGTT 1302  
Db 261 GlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPhe 280  
Qy 1303 CTCACAAATGTTGCTTCTCTGTTTACATGCTGCTAAGCAACCAACGAGTGACAGGCT 1362  
Db 281 LeuAsnAsnValAlaSerProValTyrIleAlaAlaGluGlnProThrAsnGlyGlnAla 300  
Qy 1363 TCTAATACGAGTAATAATTTACGAGATGGAGAGCTATCTCTGTAAGAATGGTGGCAA 1422  
Db 301 SerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGln 320  
Qy 1423 ---GCAGGATCCAATACTCTGGATCAGTTTCTTTGATGGAGAGGAGTAGTTTCTTTT 1479  
Db 321 AlaAlaGlySerAsnAsnSerGlySerValSerPheAspGlyGlyGlyValPhePhe 340  
Qy 1480 AGTAGCAATGTAGCTCTGGAAAGGGGAGCTATTATTATGCCAAAAGCTCTCGTGTCT 1539  
Db 341 SerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAla 360  
Qy 1540 AACTGTGGCCGTGACAAATTTTAAAGGAATATCGCTAAATGATGGTGGAGCGATTTATTA 1599  
Db 361 AsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsnAspGlyGlyAlaIleTyrLeu 380

Qy 1600 GGAGAACTCGAGAGCTCAGTTTATCTGCTGATATTCGAGATATATTTTCGATGGGAAT 1659  
Db 381 GlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsn 400  
Qy 1660 CTTAAAAGAACAGCAAGAGAAATGCTGCCGATGTTAATGCGTAACGTGTCTCTCACAA 1719  
Db 401 LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln 420  
Qy 1720 GCCATTTTCGATGGGATCGGAGGAGAAATAACGACATTAAGAGCTTAAAGCGGCGATCAG 1779  
Db 421 AlaIleSerMetGlySerGlyGlyIleThrThrLeuArgAlaLysAlaGlyHisGln 440  
Qy 1780 ATTCTCTTTAATGATCCCATCGAGATGCCAAACGGAATAACCGACGAGCGAGCTTCCC 1839  
Db 441 IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer 460  
Qy 1840 AAACCTTCTAAAATTTAAACGATGGTGAAGGATACACAGGGGATATGTTTTTGTCTAATGA 1899  
Db 461 GluProLeuLysIleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGly 480  
Qy 1900 AGCAGTACTTTGTACCAAAATGTTACGATAGACAGGAAGGATGTTTCTTGTGTAAG 1959  
Db 481 AsnSerThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLys 500  
Qy 1960 GCAAAATTTATCAGTGAATCTCTAAGTCAGACAGCTGGGAGTCTGTATATGGAAGCTGG 2019  
Db 501 AlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGly 520  
Qy 2020 AGTACATCGGATTTTGTAACTCCACACACACACAGCTCTCTGCGCTAAATCAGTTG 2079  
Db 521 SerThrLeuAspPheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeu 540  
Qy 2080 ATCAGCTTTTCCAATCTGCATTTGTCTTTCTTTTGTAGCAAAACAATGCAGTTACG 2139  
Db 541 IleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThr 560  
Qy 2140 AATCTCTCTCAATCTCCAGCGCAAGATTCATCTCTCGCAGTCATTTGGTAGCAACT 2199  
Db 561 AsnProThrAsnProAlaGlnAspSerHisProAlaIleIleGlySerThrThr 580  
Qy 2200 GCTGGTCTGTTTACAAATTAGTGGGCTATCTTTTGTAGGATTTGGATGATACAGCTTAT 2259  
Db 581 AlaGlySerValThrIleSerGlyProIlePhePheGluAspLeuAspThrAlaTyr 600  
Qy 2260 GATAGTATGATGGCTAGTTCTAATCAAAAATCAATGTCTCTCAAAATACAGTTAGGG 2319  
Db 601 AspArgTyrAspTyrLeuGlySerAsnGlnLysIleAspValLeuLysLeuGlnLeuGly 620  
Qy 2320 ACTAAGCCCCGACTAATGCCCCCATCAGATTGACTCTAGGGAATGAGATGCTTAAGTAT 2379  
Db 621 ThrGlnProSerAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyr 640  
Qy 2380 GGTATCAAGGAGCTGGAAGCTTCGGTGGGATCTCTAATACAGCAAAATAATGGTCTTAT 2439  
Db 641 GlyTyrGlnGlySerTyrPheLysLeuAlaTyrPheProAsnThrAlaAsnAsnGlyProTyr 660  
Qy 2440 ACTCTGAAGCTACATGGAATAACTGGG 2469  
Db 661 ThrLeuLysAlaThrTyrThrLysThrGly 670

## RESULT 28

US-11-109-468-169  
; Sequence 169, Application US/11109468  
; Publication No. US20050232941A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Guderian, Jeff  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Maisonneuve, Jean-Francois L.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT  
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.515C4

; CURRENT APPLICATION NUMBER: US/11/109,468  
; CURRENT FILING DATE: 2005-04-19  
; PRIOR APPLICATION NUMBER: US 10/197,220  
; PRIOR FILING DATE: 2002-07-15  
; PRIOR APPLICATION NUMBER: US 10/007,693  
; PRIOR FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: US 10/012,256  
; PRIOR FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 09/841,260  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/219,752  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: US 60/198,853  
; PRIOR FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 175  
; SEQ ID NO 169  
; LENGTH: 670  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-11-109-468-169

## Alignment Scores:

Pred. No.: 670  
Score: 3336.50 Matches: 651  
Percent Similarity: 98.8% Conservative: 11  
Best Local Similarity: 97.2% Mismatches: 7  
Query Match: 42.3% Indels: 1  
DB: Gaps: 1

US-10-701-844-1 (1-4435) x US-11-109-468-169 (1-670)

```
QY 463 GCAGAAATCATGTTCTCTCAGGAATTTAGTAGGGAGCGTTAACTGTATCATTTCCC 522
DB 1 AlaGluileMetileProGlnGlyIleTyAspGlyGluThrLeuThrValSerPhePro 20
QY 523 TATACGTGTATAGGACATCGAGTGGACACTACTGTTTTTCTCGAGGAGATTAAACATTA 582
DB 21 TyThrValIleGlyAspProSerGlyThrValPheSerIleGlyGluLeuThrLeu 40
QY 583 AAAAACTTTGACAAATCTTATGCGCTTTGCGCTTTTAAAGTTGTTTGGAACTATTAGGG 642
DB 41 LysAsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGly 60
QY 643 AGTTTTACTGTTTGGGAGAGACACTGTTGCACTTTTCGAGAACATACGGACTTCTACA 702
DB 61 SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr 80
QY 703 AATGGGGCAGCTCTAAGTAATAGCTGCTGCTGAGTGTCTTACTATTGAGGGTTTAA 762
DB 81 AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys 100
QY 763 GAATTATCTTTTCCAAATTCGAATTCATTAATCTTCCGCTACTGCTGCTCAACGACTAAT 822
DB 101 GluLeuSerPheSerAsnSerLeuLeuAlaValLeuProAlaThrThrAsn 120
QY 823 AAGGGTAGCAGACTCCGACGACAACTACACCGTCTAATGGTACTATTATTTCTAAA 882
DB 121 AsnGlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLys 140
QY 883 ACAGATCTTTGTACTCAATATGAGAGTTCTCATCTTATAGTATTTAGTCTCTGGA 942
DB 141 ThrAspLeuLeuLeuLeuAsnGlyPheSerPheTySerAsnLeuValSerGly 160
QY 943 GATGGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTTCAAGGAATTAAGAGCTTTGTGTC 1002
DB 161 AspGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCyVal 180
QY 1003 TTCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTTGTCAAGTAGTCAAGTTTCTCT 1062
DB 181 PheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrThrSerPheSer 200
QY 1063 GCTATGCTAACGAGGCTCTATTGCTTTGTAGCGAATGTTGAGGAGTAAGAGGGGA 1122
```

```
DB 201 AlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsnValAlaGlyValArgGlyGly 220
QY 1123 GGGATTGCTGTTGTTTCCAGGATGGGAGGAGGTGTCATCTATCTTCAACAGAGAT 1182
DB 221 GlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerThrSerThrGluAsp 240
QY 1183 CCAGTAGTAAGTTTTCAGAAATCTACCGGTAGAGTTTGTAGTGGAAAGTAGCCGAGTA 1242
DB 241 ProValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgVal 260
QY 1243 GGAGGAGGATTTACTCTTACGGGAACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1302
DB 261 GlyGlyGlyIleTySerTyGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPhe 280
QY 1303 CTCACAAATGTTGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1362
DB 281 LeuAsnAsnValAlaSerProValTyIleAlaGluGlnProThrAsnGlyGlnAla 300
QY 1363 TCTAATACGAGTAATAATACGAGATGGAGAGGTATCTTCTGTGTAGAGTGGTGGCAA 1422
DB 301 SerAsnThrSerAspAsnTyGlyAspGlyAlaIlePheCysLysAsnGlyAlaGln 320
QY 1423 ---GCAGATCCAATAACTCTCGATCAGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1479
DB 321 AlaAlaGlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePhe 340
QY 1480 AGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTATGCCAAAGAGCTCTCGGTGCT 1539
DB 341 SerSerAsnValAlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAla 360
QY 1540 AACTGTGGCCCTGTACAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGATTATTATA 1599
DB 361 AsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsnAspGlyGlyAlaIleTyLeu 380
QY 1600 GGAGAACTCTGGAGAGCTCAGTTTCTGCTGATTTATGGAGATATTATTTTCATGCGAAT 1659
DB 381 GlyLysSerGlyGluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsn 400
QY 1660 CTTAAAGAACACCAAGAGAAATGCTGCCGATGTTAATGGCGTAACTGTGCTCTCACAA 1719
DB 401 LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln 420
QY 1720 GCATTTTCGATGGGATCGGGAGGAAAATAACACATTAAGAGCTAAAGCAGGCGATCAG 1779
DB 421 AlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGln 440
QY 1780 ATTCTTTTAAATGATCCCATCGAGATGGCAACCGGAATAACACGAGCCGAGCTTCTCC 1839
DB 441 IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer 460
QY 1840 AAACCTCTAAAATTAACGATGCTGAAGGATACACAGGGGATATTGTTTTGCTAATGGA 1899
DB 461 GluProLeuLysIleAsnAspGlyGluGlyTyThrGlyAspIleValPheAlaAsnGly 480
QY 1900 AGCAGTACTTTGTACCAAAATTTACGATAGACAGGAGGAGGATTTCTTCTGTGAAAAG 1959
DB 481 AsnSerThrLeuTyGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLys 500
QY 1960 GCAAAAATTATCAGTGAATTTCTTAAGTCAGACAGGTGGGAGTGTGTATATGGAAGCTGG 2019
DB 501 AlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyMetGluAlaGly 520
QY 2020 AGTACATGGATTTTGTAACTCCACACCAACCAACGCTCTCCGCTAATCAGTTG 2079
DB 521 SerThrLeuAspPheValThrProGlnProProGlnGlnProProAlaAlaAsnGlnLeu 540
QY 2080 ATCACGCTTTCCAATCTGCAATTTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2139
DB 541 IleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAlaValThr 560
QY 2140 AATCCTCTCAACCAATCTCCAGCGCAAGATTCTATCTCTCAGTCAATTTGTAGACAACT 2199
DB 561 AsnProProThrAsnProAlaGlnAspSerHisProAlaIleIleGlySerThrThr 580
```

```
QY 2200 CTTGCTGTTTCAATAGTGGGCTATCTTTTGTAGGATTTGGATGATACAGCTTAT 2259
Db 581 AlaGlySerValThrIleSerGlyProIlePhePheGluAsePLeuAsePThrAlaTy 600
QY 2260 GATAGGTATGATGGCTAGGTTCTTAATCAAAAATCAATGTCCTGAAATACAGTTAGG 2319
Db 601 AspArgTyAsePTrpLeuGlySerAsnGlnLysIleAsePValLeuLysLeuGlnLeuGly 620
QY 2320 ACTAAGCCCCAGCTAAATGCCCATCAGATTGACTTAGGGAATGAGATGCCTAAGTAT 2379
Db 621 ThrGlnProSerAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTy 640
QY 2380 GGCTATCAAGGAAGCTGAAGCTGCGTGGGATCTTAATACAGCAAAATAATGGTCTTAT 2439
Db 641 GlyTyGlnGlySerTrpLysLeuAlaTrpAsePProAsePThrAlaAsnAseGlyProTy 660
QY 2440 ACTCTGAAGCTACATGGACTAAAACCTGGG 2469
Db 661 ThrLeuLysAlaThrTrpThrLysThrGly 670

RESULT 29
US-10-701-844-17
; Sequence 17, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-399
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-17

Alignment Scores:
Pred. No.: 6,71e-218 Length: 505
Score: 2552.00 Matches: 505
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.4% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-701-844-17 (1-505)
QY 466 GAAATCATGTTCTCTCAAGGAATTTACGATGGGAGAGCTTAACCTGTATCATTTCCCTAT 525
Db 1 GluIleMetValProGlnGlyIleTyAsePLeuThrValSerPheProTy 20
QY 526 ACTGTTATAGGATCCGAGTGGGACTACTGTTTTCGACAGGAGATTACATTAAA 585
Db 21 ThrValIleGlyAsePProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
QY 586 AATCTTGACAAATCTATTGACCTTGCCTTTAAAGTTGTTTGGGAACCTTATTAGGAGT 645
Db 41 AsnLeuAsePAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuGlySer 60
QY 646 TTTACTGTTTATAGGAGAGGACACTGTTGACTTTTCGAGAAACATACGAGCTTCTCAAAAT 705
Db 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleAsePThrSerThrAse 80
QY 706 GGGGCGAGCTTAAGTAATAGCCCTGATGACTGTTTACTATTGAGGGTTTAAAGAA 765
Db 81 GlyAlaAlaLeuSerAsnSerAlaAlaAsePLeuPheThrIleGlyPheLysGlu 100
```

```
QY 766 TTATCTCTTTTCCAAATTCGAATTCATTACTTGCCCTACTGCTGCTGCAACGACTAATAAG 825
Db 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
QY 826 GGTAGCCAGACTCCGACGACAACTACACCGCTCAATGGTACTATTATTCTTAAACA 885
Db 121 GlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLysThr 140
QY 886 GATCTTTTGTACTCAATAATGAGAAGTTCTATTCTATAGTAATTTAGTCTCTGGAGAT 945
Db 141 AspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTySerAsnLeuValSerGlyAseP 160
QY 946 GGGGAGGCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTC 1005
Db 161 GlyGlyAlaIleAsePAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
QY 1006 CAAGAAATACCTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCT 1065
Db 181 GlnGluAsnThrAlaGlnAlaAsePglyLysCysGlnValValThrSerPheSerAla 200
QY 1066 ATGGCTAACGAGGCTCTCTATTGCTCTTTGTAGCGAATGTTGCAGGAGTAAGAGGGAGGG 1125
Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
QY 1126 ATTGCTGCTGTTCAAGGATGGGAGGAGTGTCATCATCTACTTCAACAGAGATCCA 1185
Db 221 IleAlaAlaValGlnAsePglyGlnGlnGlyValSerSerSerThrSerThrGluAsePPro 240
QY 1186 GTAGTAAAGTTTTCGAGAAATACCTGCGTACAGTTTGTATGGGAGCTAGCCGAGTAGGA 1245
Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAsePglyAsnValAlaAargValGly 260
QY 1246 GGAGGATTTACTCTCTACGGCAAGCTGCTTCTCTGAATAATGGAACCTTGTCTCTC 1305
Db 261 GlyGlyIleTySerTyGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280
QY 1306 AACAAATGTTCTCTCTGTTTACATGCTGCTAGCAACCAACAAAGTGGAGAGCTTCT 1365
Db 281 AsnAsnValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
QY 1366 AATACGAGTAATAATACGGAGATGGAGGAGCTATCTCTGTAAGAAATGGTCGCAAGCA 1425
Db 301 AsnThrSerAsnAsnTyGlyAsePglyLysAlaIlePheCysLysAsnGlyAlaGlnAla 320
QY 1426 GGATCCATAACTCTGATCAAGTTTCTCTTGTATGGAGGAGGAGTAGTTTCTTTAGTAGC 1485
Db 321 GlySerAsnAsnSerGlySerValSerPheAsePglyGluGlyValValPhePheSerSer 340
QY 1486 AATGTAGCTGCTGGGAAAGGGGAGCTATTTATGCCAAAAGCTCTCGGTTGCTAACTGT 1545
Db 341 AsnValAlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCys 360
QY 1546 GGCCCTGTACAAATTTTAAAGGAATATCCGTAATGATGGTGGAGCGATTTATTAGAGAA 1605
Db 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAsePglyAlaIleTyLeuGlyGlu 380
QY 1606 TCTGAGAGCTCAGTTTATCTGCTATATGGAGATATTTTTCGATGGGAATCTTAAA 1665
Db 381 SerGlyGluLeuSerLeuSerAlaAsePtyGlyAsePillePheAsePglyAsnLeuLys 400
QY 1666 AGAACAGCAAGAGAGATGCTGCCGATGTTAATGGCGTAACCTGTCTCTCACAGCCATT 1725
Db 401 ArgThrAlaLysGluAsnAlaAlaAsePValAseGlyValThrValSerSerGlnAlaIle 420
QY 1726 TCGATGGGATCGGGAGGAGAAATACGACTTAAGAGCTAAAGCAGGAGGATCAGATCTTC 1785
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
QY 1786 TTTAATGATCCCATCGAGATGGCAACCGAAATTAACCGACCGAGCGAGCTTCTTCAAACTT 1845
Db 441 PheAsnAsePProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460
```



## RESULT 31

US-10-931-779-17  
 ; Sequence 17, Application US/10931779  
 ; Publication No. US20050048557A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; APPLICANT: Pace, John  
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
 ; FILE REFERENCE: BP104  
 ; CURRENT APPLICATION NUMBER: US/10/931,779  
 ; CURRENT FILING DATE: 2004-09-01  
 ; PRIOR APPLICATION NUMBER: 09/542,520  
 ; PRIOR FILING DATE: 2000-04-03  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 17  
 ; LENGTH: 505  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia sp.  
 US-10-931-779-17

Alignment Scores:  
 Pred. No.: 6 71e-218 Length: 505  
 Score: 2552.00 Matches: 505  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 32.4% Indels: 0  
 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-931-779-17 (1-505)

QY 466 GAAATCATGTTCTCAAGGAATTTACGATGGGAGAGCTTAATGTATCATTTCCCTAT 525  
 Db 1 GluileMetValProGlnGlyIleTyraepGlyGluThrLeuThrValSerPheProTy 20  
 QY 526 ACTGTTATAGGAGATCCGATGGGACTACTGTTTCTTCGAGGAGAGTTAACTAA 585  
 Db 21 ThrValIleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLys 40  
 QY 586 AATCTTGCAATTTCTATTCAGCTTTCCTTAAAGTTGTTTGGGAATTTATTAGGAGT 645  
 Db 41 AsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60  
 QY 646 TTTACTGTTTATAGGAGAGACACTCGTTGACTTTTCGAGAACATACGACTTCTCAAA 705  
 Db 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80  
 QY 706 GGGGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGTTTTAAGNA 765  
 Db 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100  
 QY 766 TTATCTCTTTTCCAAATTCATTACTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825  
 Db 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120  
 QY 826 GTAGCCAGACTCCGACGACATCTACACCGTCTAATGGTACTATTATTCTTAAACA 885  
 Db 121 GlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLysThr 140  
 QY 886 GATCTTTTGTACTCAATTAATAGAGATTCTCATCTATAGTAATTTAGTCTCTGAGAT 945  
 Db 141 AspLeuLeuLeuAsnAsnGlnLysPheSerPheTySerAsnLeuValSerGlyAsp 160  
 QY 946 GGGGAGCTATAGTCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTCTC 1005  
 Db 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180  
 QY 1006 CAAGAAATPACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCAACGATTTCTGCT 1065  
 Db 181 GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSerAla 200  
 QY 1066 ATGGCTAACGAGCTCTATTGCTTTGTAGCGAATGTTGCGAGGATAGAGGGGAGGG 1125

Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220  
 QY 1126 ATTCTGCTGTTAGGATGGCAGCAGGAGTGTCACTACTTCAACAGAGATCCA 1185  
 Db 221 IleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspPro 240  
 QY 1186 GTAGTAAGTTTTTCCAGAAATACCTCGGTAGAGTTTGTATGGACGTAGCCGAGTAGGA 1245  
 Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260  
 QY 1246 GGAGGATTTTACTCTACGGGAACGTTGCTTCTGAATAATGAAAAACCTTGTTCCTC 1305  
 Db 261 GlyGlyIleTySerTyGlyAsnValAlaPheLeuAsnGlnGlyLysThrLeuPheLeu 280  
 QY 1306 AACAAATGTTGCTTCTCTGTTTACATTGCTCTAAGCAACCAACAAAGTGGAGCTTCT 1365  
 Db 281 AsnAsnValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300  
 QY 1366 AATACGAGTAATAATTACGGAGATGGAGGAGCTATCTCTTAAGATGGTGGCAAGCA 1425  
 Db 301 AsnThrSerAsnAsnTyGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320  
 QY 1426 GGATCCAATACTCTGGATCAGTTTCTTTCATCGAGAGGAGTAGTTTCTTTAGTAGC 1485  
 Db 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340  
 QY 1486 AATGTAGCTGCTGGGAAAGGGGAGCTATTATATCCAAAAAGCTCTCGTTGCTAACTGT 1545  
 Db 341 AsnValAlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCys 360  
 QY 1546 GGCCCTGTACAAATTTTAAAGGAATATCGCTAATGATGCTGGAGGAGTTATTAGGAAA 1605  
 Db 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyLeuGlyGlu 380  
 QY 1606 TCTGAGAGCTCAGTTTATCTGCTGATATGAGAGATATTATTTTCGATGGGAATCTTAAA 1665  
 Db 381 SerGlyGluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsnLeuLys 400  
 QY 1666 AGAACAGCCAAAGAGAATGCTGCCGATTTAATGGCTAACTGTGCTCTCAAGCCATT 1725  
 Db 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420  
 QY 1726 TCGATGGATCGGAGGGGAAAAATAACGACATTAGAGCTTAACGCGGCGATCAGATTCTC 1785  
 Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440  
 QY 1786 TTTAATGATCCCATCGAGATGCAACCGAAATAACCGACCGAGCGAGCTTTCCAAACCTT 1845  
 Db 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460  
 QY 1846 CTAATAATTAACGATGGTGAAGGATACACAGGGGATATTGTTTTGCTTAATGGAAGCAGT 1905  
 Db 461 LeuLysIleAsnAspGlyGlyTyThrGlyAspIleValPheAlaAsnGlySerSer 480  
 QY 1906 ACTTTGTACCAAAATGTTACGATAGCAAGGAGGATGTTCTTCGTGAAAAGGCAAAA 1965  
 Db 481 ThrLeuTyGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500  
 QY 1966 TTATCAGTGAATTCCT 1980  
 Db 501 LeuSerValAsnSer 505

## RESULT 32

US-10-701-844-36  
 ; Sequence 36, Application US/10701844  
 ; Publication No. US20040067524A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; APPLICANT: Pace, John  
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
 ; FILE REFERENCE: 7969-086-999  
 ; CURRENT APPLICATION NUMBER: US/10/701,844





```
QY 1030 GGGGAGCTTCTCAAGTAGTCACCAAGTTCTCTGCTATGGCTAACGAGGCTCCTATTGCC 1089
Db 1 GlyGlyAlaCysGlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAla 20
QY 1090 TTTGTACGGAATGTTGAGAGTAAGAGGGGAGGAGTTCCTGCTTCCAGATGGGCAG 1149
Db 21 PheValAlaAsnValAlaGlyValAlaGlyGlyIleAlaValAlaGlnAspGlyGln 40
QY 1150 CAGGAGTGTCATCATCTTCTCAACAGAGATCCAGTAGTAGTATTTTTCAGAAATACT 1209
Db 41 GlnGlyValSerSerThrSerThrGluAspProValValSerPheSerArgAsnThr 60
QY 1210 GCGGTAGAGTTTGTAGTGGAAAGTACGCCGAGTAGGAGGAGGATTTACTCTCAGGGAAC 1269
Db 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsn 80
QY 1270 GTTCTCTTCCGAAATAATGAAAAACCTTGTCTCAACAATGTGCTTCTCCTGTTTAC 1329
Db 81 ValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyr 100
QY 1330 ATTCTGCTAAGCAACCAAGTGGAGAGGCTTCTTAATACGAGTAATAATTACGGAGAT 1389
Db 101 IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAsp 120
QY 1390 GGAGGAGCTATCTTCTGTAGAATGGTGGCAAGCAGGATCCAAATAACTCTGGATCAGTT 1449
Db 121 GlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal 140
QY 1450 TCCTTTGTAGAGAGGAGTAGTATTTCTTTAGTAGCAATAGTAGCTGGGAAAGGGGA 1509
Db 141 SerPheAspGlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGly 160
QY 1510 GCTATTATGCGAAAAAGCTCCGGTGTCTTAACGTGGCCCTGTACAAATTTTAAAGGAAT 1569
Db 161 AlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn 180
QY 1570 ATCGCTAATGATGGTGGAGCAATTTATTAGAGAACTTGGAGAGCTCAGTTTATCTGCT 1629
Db 181 IleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAla 200
QY 1630 GATTATGAGATATTTATTTTCATGGGAATCTTAAAGAACGCCAAAGAGATGCTGCC 1689
Db 201 AspTyrGlyAspIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAla 220
QY 1690 GATGTTAATGGGTAACTGTCTCTCAAGCCATTTGATGGGATCGGAGGGAATA 1749
Db 221 AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIle 240
QY 1750 ACGACATTTAAGACTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGCA 1809
Db 241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAla 260
QY 1810 AACGAAATAACACCGACGAGCTCTCCAAACTTCTAAATAATTAACATGCTGAAGA 1869
Db 261 AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGly 280
QY 1870 TACACAGGGATTTGTTTCTTAATGAACAGTACTTTCTACAAATGCTTACGATA 1929
Db 281 TyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIle 300
QY 1930 GAGCAAGGAAGATTGTTCTTCGTGAAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAG 1989
Db 301 GluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGln 320
QY 1990 ACAGTGGGAGTCTGTATATGAAGCTGGAGTACATGGGATTTGTAACTCCACACCA 2049
Db 321 ThrGlyGlySerLeuTyrMetGluAlaGlySerThrTrpAspPheValThrProGlnPro 340
QY 2050 CCACACAGCTTCTGCGCTAATCAGTTGATCAGCTTCCAACTCCAACTGCAATTTGCTCTT 2109
Db 341 ProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu 360
QY 2110 TCTTCTTTGTAGCAACAAATGCAAGTTACGAATCTCTCTTACCAATCTCCAGCGCAAGAT 2169
```

```
Db 361 SerSerLeuLeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAsp 380
QY 2170 TCTCATCTCAGTCAGTTCAGTTCAGCAACTGCTGTTCTGTTACAATTTAGTGGGCTTATC 2229
Db 381 SerHisProAlaValIleGlySerThrAlaGlySerValThrIleSerGlyProIle 400
QY 2230 TTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAATCAA 2289
Db 401 PhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGln 420
QY 2290 AAAATCAATCTCTGAAATTTACAGTTAGGACTTAAGCCCCCAGCTAATGCCCATCAGAT 2349
Db 421 LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAsp 440
QY 2350 TTGACTCTACGGAATGAGATGCTTACGCTATGAGTATGGCTATCAAGGAAGCTCGAAGCTT 2403
Db 441 LeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeu 458

RESULT 34
US-10-931-779-36
; Sequence 36, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: B104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-36

Alignment Scores:
Pred. No.: 6.52e-200 Length: 458
Score: 2350.00 Matches: 458
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.8% Indels: 0
DB: Gaps: 0

US-10-701-844-1 (1-4435) x US-10-931-779-36 (1-458)
QY 1030 GGGGAGCTTCTCAAGTAGTCACCAAGTTTCTCTGCTATGGCTAACGAGGCTCCTATTGCC 1089
Db 1 GlyGlyAlaCysGlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAla 20
QY 1090 TTTGTACGGAATGTTGAGAGTAAGAGGGGAGGAGTTCCTGCTTCCAGATGGGCAG 1149
Db 21 PheValAlaAsnValAlaGlyValAlaGlyGlyIleAlaValAlaGlnAspGlyGln 40
QY 1150 CAGGAGTGTCATCATCTTCTCAACAGAGATCCAGTAGTAGTATTTTTCAGAAATACT 1209
Db 41 GlnGlyValSerSerThrSerThrGluAspProValValSerPheSerArgAsnThr 60
QY 1210 GCGGTAGAGTTTGTAGTGGAAAGTACGCCGAGTAGGAGGAGGATTTACTCTCAGGGAAC 1269
Db 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsn 80
QY 1270 GTTCTCTTCCGAAATAATGAAAAACCTTGTCTCAACAATGTGCTTCTCCTGTTTAC 1329
Db 81 ValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyr 100
QY 1330 ATTCTGCTAAGCAACCAAGTGGAGAGGCTTCTTAATACGAGTAATAATTACCGAGAT 1389
Db 101 IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAsp 120
```

```

1390 GGAGAGCTATCTTCTGTAAGAATCGTGGCGCAAGCAGGATCCAAATAACTCTGATCAGTT 1449
121 GlyGlyAlaIlePheCysIysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal 140
1450 TCCTTTGATGGAGAGGAGTGTCTTTCTTAGTAGCAATGTAGCTGCTGGGAAAGGGGA 1509
141 SerPheAspGlyGlyValValPhePheSerSerAsnValAlaAlaGlyGlyGly 160
1510 GCTATTTATGCCAAAGCTCTCGGTTGTAACTGTGGCCCTGTACAAATTTTAGGAAT 1569
161 AlaIleTyrAlaIysIysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn 180
1570 ATCCGCTAATGATGGCGGATTTATTTAGAGATCTGGAGAGCTCAGATTTATCTGCT 1629
181 IleAlaAsnAspGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAla 200
1630 GATTATGGAGATATTATTTTCGATGGAAATCTTAAAGAACACAGCCAAAGAGAAATGCTGCC 1689
201 AspTyrGlyAspIleIlePheAspGlyAsnLeuLeuYrArgThrAlaIysGluAsnAla 220
1690 GATGTTAATGGGTAATCTGTCTCTCAAGCCATTTTCGATGGGATCGGAGAGGAAATA 1749
221 AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIle 240
1750 ACACATTAAGAGCTAAAGCAGGCGATCATCTCTTTAATGATCCCATCGAGATGGCA 1809
241 ThrThrLeuArgAlaIysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAla 260
1810 AACGGAATAACAGCAGCGAGTCTTCCAAACTTCTAAATAATTAACGATGTGAAGGA 1869
261 AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGly 280
1870 TACACAGGGGATATGTTTTGCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATA 1929
281 TyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIle 300
1930 GAGCAAGGAAGATGTTCTTCTGTAAGAGGCAAAATTTATCAGTGAATCTCTAAGTCAG 1989
301 GluGlnGlyArgIleValLeuArgGlyAlaLysLeuSerValAsnSerLeuSerGln 320
1990 ACAGTGGGAGCTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACACCA 2049
321 ThrGlyGlySerLeuTyrMetGluAlaGlySerThrThrPheValThrProGlnPro 340
2050 CCACACAGCTCTCTCGCGTAACTAGTTGATCACGCTTTCCAACTCGCATTTGTCTCT 2109
341 ProGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu 360
2110 TCTCTTTGTTAGCAACATGAGTTACGATTAAGATCTCTTACCAATCTCCGCGCAGAT 2169
361 SerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAsp 380
2170 TCTCATCTCAGCTCATTGTGTAGCACAACTGCTGTTCTGTTCAATTTAGTGGCCTATC 2229
381 SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIle 400
2230 TTTTCTTGGAGATTTGATGATACAGCTTATGATAGGTATGATTTGGCTAGGTTCTAATCAA 2289
401 PhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGln 420
2290 AAAATCAATGTCCTGAAATTTACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCATGAT 2349
421 LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProProAlaAlaProSerAsp 440
2350 TTGACTCTAGGATGAGTCCCTAAGTATGGCTATCAAGGAGCTGGAAGCTT 2403
441 LeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeu 458

```

RESULT 35

US-10-701-844-37

; Sequence 37, Application US/10701844

; Publication No. US20040067524A1

```

; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Page, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Chlamydia sp.
; US-10-701-844-37

Alignment Scores:
Pred. No.: 37e-145 Length: 325
Score: 1735.00 Matches: 325
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.0% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-701-844-37 (1-325)

```

```

QY 2443 CTGAAAGCTACATGAGCTAAATCTGGTATATAATCTGGGCTGAGGAGTAGCTTCTTTG 2502
DB 1 LeuLysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeu 20
QY 2503 GTTCCAATAGTTATGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATTCAA 2562
DB 21 ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln 40
QY 2563 GCAAGTGTGGATGGGCGCTCTTATTGTCCGAGATATTATGGTTCTCGAATTTTCGAATTC 2622
DB 41 AlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhe 60
QY 2623 TTCTATCATGACCGCGATCTTTAGGTCAGGATATCGGTATATATAGTGGGTTTATTC 2682
DB 61 PheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSer 80
QY 2683 TTAGGAGCAAACTCTTACTTTTGGATCATCGATTTGTTGTTAGCTTATCCGAGATTTT 2742
DB 81 LeuGlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPhe 100
QY 2743 GGTAGATCTAAAGATTATGTAGTGTGCTTCCATCAATCATCATCTTGCATAGGATCCGTT 2802
DB 101 GlyArgSerLysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerVal 120
QY 2803 TATCTATCTACCCAAACAGCTTTTATGTGATCTTATTTGTCGAGATGCGGTTTATCCGT 2862
DB 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArg 140
QY 2863 GCTAGCTAGCGTTTGGGATTCAGCATATGAAACCTCATATACATTTTCAGAGGAGAGC 2922
DB 141 AlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSer 160
QY 2923 GATGTTCTGGGATTAATCTCTGCTGGCTGGAGAGATGGAGCGGATTTACCGATTGTG 2982
DB 161 AspValArgTrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleVal 180
QY 2983 ATTACTCCATCTAAGCTCTTATTTGAATGAGTTCCGCTCTTCTTCGTCAGCTGAGTTTCT 3042
DB 181 IleThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSer 200
QY 3043 TATGCGCATCATCAATCTTTTACAGAGGAGGCGATCAAGCTCGGCGATTCAGAGGAGGA 3102
DB 201 TyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGly 220

```

```
QY 3103 CATCTCTAAATCTATCAGTTCCCTGCTGGAGTGAAGTTTGATCGATGTTCTAGTACACAT 3162
Db 221 HisLeuLeuAenLeuSerValProValGlyValValPheAspArgCysSerSerThrHis 240
QY 3163 CCTAATAATATAGCTTTATGCGGCTTATATCTGATGCTGATCGACCATCTCTGGT 3222
Db 241 ProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGly 260
QY 3223 ACTGAGACACGCTCTATCCATCCATCAAGACATGACACACAGATGCTCTTTCATTAGCA 3282
Db 261 ThrGluThrThrLeuLeuSerHisGlnGlnThrThrAspAlaPheHisLeuAla 280
QY 3283 AGACATGAGTTGGTTAGAGGATCTATGATGCTTCTTAACAGTAATATAGAGTA 3342
Db 281 ArgHisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluVal 300
QY 3343 TATGCCCATGGAAGATATGATGATCGATGCTCTCGAGGCTATGGTTTGATGTCAGGA 3402
Db 301 TyrGlyHisGlyValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluVal 320
QY 3403 AGTAGAGTCCGGTTC 3417
Db 321 SerArgValArgPhe 325
RESULT 36
US-10-766-711-37
; Sequence 37, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-37
Alignment Scores:
Pred. No.: 3 7e-145 Length: 325
Score: 1735.00 Matches: 325
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.0% Indels: 0
DB: 4 Gaps: 0
US-10-701-844-1 (1-4435) x US-10-766-711-37 (1-325)
QY 2443 CTGAAGCTACATGGAATAAACTGGGTATAATCTGGCCCTGAGCGAGTAGCTCTTTG 2502
Db 1 LeuLysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeu 20
QY 2503 GTTCCAAATAGTTATGGGATCCATTTAGATATACGATCGCGCATTCAGCAATTCAA 2562
Db 21 ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln 40
QY 2563 GCAAGTGTGATGGGCGCTTATGTCGAGGATATGAGTTTCTGAGTTTCGAAATTC 2622
Db 41 AlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhe 60
QY 2623 TTCTATCATGACCGCATGCTTTAGGTCAGGATATCGGTATATATAGTGGGGTATATCC 2682
Db 61 PheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyTyrSer 80
QY 2683 TTAGGAGCAACTCTTACTTTCGATCATCGATGTTTGGTCTAGCATTTACCGAGTATTT 2742
```

```
Db 81 LeuGlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPhe 100
QY 2743 GGTAGATCTAAGAAATATAGTAGTGTGCTGTTCCAAATCATCATGCTTGATAGGATCCGTT 2802
Db 101 GlyArgSerLysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerVal 120
QY 2803 TATCTATCTACCCCAACAGCTTTATGTCGATCCCTATTGTTGCGAGATGCTTTATCCGT 2862
Db 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArg 140
QY 2863 GCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATATTTCAGAGAGAGAGC 2922
Db 141 AlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSer 160
QY 2923 GATGTTCTGCGGATAAATACTGCTGCTCGAGAGATTTGAGCGGGATTTACCGATTGTG 2982
Db 161 AspValArgTyrAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleVal 180
QY 2983 ATTACTCCATCTAAGCTCTATTGTAATGAGTTGGTCTCTTTCGTCGAAGCTGAGTTTCT 3042
Db 181 IleThrProSerLysLeuTyrLeuAsnIleLeuArgProPheValGlnAlaGluPheSer 200
QY 3043 TATGCCGATCATGAATCTTTTACAGAGAGAGCGCATCAAGCTCGGGCATTTCAAGAGCGGA 3102
Db 201 TyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGly 220
QY 3103 CATCTCTAAATCTATCAGTTCCCTGTTGGAGTGAAGTTTGATGCGATGTTCTAGTACACAT 3162
Db 221 HisLeuLeuAenLeuSerValProValGlyValValPheAspArgCysSerSerThrHis 240
QY 3163 CCTAATAATATAGCTTTATGCGGCTTATATCTGATGCTGATCGACCATCTCTGGT 3222
Db 241 ProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGly 260
QY 3223 ACTGAGACACGCTCTATCCATCCATCAAGACATGACACACAGATGCTCTTTCATTAGCA 3282
Db 261 ThrGluThrThrLeuLeuSerHisGlnGlnThrThrAspAlaPheHisLeuAla 280
QY 3283 AGACATGAGTTGGTTAGAGGATCTATGATGCTTCTTCTAAACAAGTAATATAGAGTA 3342
Db 281 ArgHisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluVal 300
QY 3343 TATGCCCATGGAAGATATGATGATCGATGCTCTCGAGGCTATGGTTTGATGTCAGGA 3402
Db 301 TyrGlyHisGlyValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluVal 320
QY 3403 AGTAGAGTCCGGTTC 3417
Db 321 SerArgValArgPhe 325
RESULT 37
US-10-931-779-37
; Sequence 37, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-37
Alignment Scores:
Pred. No.: 3 7e-145 Length: 325
```

Score: 1735.00 Matches: 325  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 22.0% Indels: 0  
DB: 5 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-931-779-37 (1-325)

QY 2443 CTGAAGCTACATGAGTAACTCGGTATATATCTGGGCCTGAGGAGTAGCTCTTTG 2502  
Db 1 LeuLysAlaThrTyrThrIleThrGlyTyrAsnProGlyProGluArgValAlaSerLeu 20  
QY 2503 GTTCCAAATAGTTATGGGATCCATTTAGATATACGATCTCGCATTCAGCAATCAA 2562  
Db 21 ValProAsnSerLeuTyrGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln 40  
QY 2563 GCAAGTGTGGGCGCTTATCTGTCGAGGATATGCGTTTCTGAGTTTCGATTTTC 2622  
Db 41 AlaSerValAspGlyArgSerTyrCysArgGlyLeuTyrValSerGlyValSerAsnPhe 60  
QY 2623 TTCTATCATGACCGCATCTTTAGGTCAGGATATCGGTATATATAGTGGGGTTATTC 2682  
Db 61 PheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSer 80  
QY 2683 TTAGGACCAACTCTTACTTTGGATCATCGATGTTTGGTCTAGCATTTACCGAATATT 2742  
Db 81 LeuGlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPhe 100  
QY 2743 GGTAGATCTAAAGATATGTAGTGTCTGTTCCAAATCATCATCTGTCATAGGATCCGTT 2802  
Db 101 GlyArgSerIleAspTyrValValCysArgSerAsnHisAlaCysIleGlySerVal 120  
QY 2803 TATCTATCTACCAACAAGCTTTATGTGGATCTTATGTTCGAGATCGGTTTATCCGT 2862  
Db 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArg 140  
QY 2863 GCTAGTACGGTTTGGGATCAGCATATCAAACTCATATACATTTGTCAGAGGAGAGC 2922  
Db 141 AlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSer 160  
QY 2923 GATGTTCTGGGATATAACTCTCTGCTGGGAGAGATTCGAGCGGATTTACCGATTGTG 2982  
Db 161 AspValArgTyrAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleVal 180  
QY 2983 ATTACTCCATCTAAGCTCTATTTGAATGAGTTCGCTCTTTCGTCAGAGTGAATTTCT 3042  
Db 181 IleThrProSerIleLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSer 200  
QY 3043 TATGCGCATCATGAATCTTTTACAGAGGAGGCGATCAAGCTCGGCGATTCAGAGCGGA 3102  
Db 201 TyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGly 220  
QY 3103 CATCTCTTAATCTATCAGTTCTCTGTTGGAGTGAAGTTTCATCGATGTTCTAGTACACAT 3162  
Db 221 HisLeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHis 240  
QY 3163 CCTAATAAATATAGCTTTATGGGCGTTATATCTGTGATGCTTATCCACCATCTCTGGT 3222  
Db 241 ProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGly 260  
QY 3223 ACTGAGACACGCTCTATCCCATCAAGACATGGAACAGATGCTTTTATTAGCA 3282  
Db 261 ThrGluThrThrLeuLeuSerHisGlnGluThrTyrThrThrAspAlaPheHisLeuAla 280  
QY 3283 AGACATCGAGTTGGTTAGGAGTCTATGTATGCTCTCTCTAAACAAGTAATATAGAAGTA 3342  
Db 281 ArgHisGlyValValValArgLysSerMetTyrAlaSerLeuThrSerAsnIleGluVal 300  
QY 3343 TATGGCCATGGAATATGATATCGAGATGCTTTCTGAGGCTATGTTGAGTGCAGGA 3402  
Db 301 TyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGly 320  
QY 3403 AGTAGATCCGGTTC 3417

Db 321 SerArgValArgPhe 325  
RESULT 38  
US-09-841-132-325  
; Sequence 325, Application US/09841132  
; Patent No. US20020061848A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: PROBST, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C8  
; CURRENT APPLICATION NUMBER: US/09/841,132  
; CURRENT FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 325  
; LENGTH: 631  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-09-841-132-325

Alignment Scores:  
Pred. No.: 3-86e-132 Length: 631  
Score: 1590.50 Matches: 316  
Percent Similarity: 88.7% Conservative: 7  
Best Local Similarity: 86.8% Mismatches: 19  
Query Match: 20.2% Indels: 22  
DB: 4 Gaps: 4

US-10-701-844-1 (1-4435) x US-09-841-132-325 (1-631)

QY 3391 TTGAGTCGAGGAGTAGATCGGTCTTAAATAATAT-----TGGTTAGATAGTTAA 3441  
Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisTyrArgPro----- 146  
QY 3442 GTGTTAGCGATCGCTTTTCTTTGAGATCTACATCATTTTGTGTTTGTGTTTGTGTGT 3501  
Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164  
QY 3502 TCCTATTCTATGAGTTCGAGCTCTCTCAAGTGTTAACGCTTAATGTAACCATCTCT 3561  
Db 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrPro 184  
QY 3562 TTTAAGGAGACGATTTTACTTTGAATGAGAGCTCGCTTTTGTCTCAATGTCTATGAGGA 3621  
Db 185 PheLysGlyAspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGly 204  
QY 3622 GCTGAAGAAGGTTTCGATTATCTCAGCTAATGGGACAATTTAAGATTACCGACAAAAC 3681  
Db 205 AlaGluAsnGlySerIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 224  
QY 3682 CATACATTATCATTTACAGATTCTCAAGGCCAGTTCTTCAAAATATTATGCTTCAATTC 3741  
Db 225 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSer 244  
QY 3742 GCAGGAGACACTTACTCTGAGAGATTTTTCAGTCTGATGTTCTCGAAAAATGTTTCT 3801  
Db 245 AlaGlyGluThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSer 264  
QY 3802 TGGCGAGAAAGGAATGATCTCCGGGAAACCGTGATTATTTCCGAGAGCAGCGAAGTG 3861  
Db 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284  
QY 3862 ATTTTCTGGGATAACTCCGTTGGGTATTCTCTTTTATCTACTGTGCTCAACCTCATCATCA 3921  
Db 285 IlePheThrAspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304  
QY 3922 ACTCCGCTCTCT----- 3945  
Db 305 ThrProProAlaProAlaProAlaProAlaSerSerLeuSerProThrValSer 324

```
QY 3946 GATGCTCGAAAGGGTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAA 4005
Db 325 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValIys 344
QY 4006 AAAGGGTCATGTCGATAAATCCCGGAATTTCCGAAACAGTTTTCGAGGTAAGAAT 4065
Db 345 LysGlyValMetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSer 364
QY 4066 AATAATAATGCTGGTGGGAGGCGAGTGGGTTCGCTACACCATCAAGTACGACTTTTAC 4125
Db 365 AsnAsnAsnAlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheTh 382
QY 4126 AGTTAAACCTGTAAGGGAAAGTTTCTTTTCACAGATAACGAGTCTTTCGGAGGCGG 4185
Db 382 rValIysAsnCysLysGlyValSerPheThrAspAsnValAlaSerCysGlyGlyG1 402
QY 4186 AGTGGTTTATAAGGCATTGCTTTTCAAGACAAATCAAGGAGGCATATCTTCCGAGG 4245
Db 402 yValValTyrLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgG1 422
QY 4246 GAACACAGCATACGATGATTTAAGGATTTCTGCTCTACTAATCAGGATCAGAATACGGA 4305
Db 422 yAsnThrAlaTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrG1 442
QY 4306 GACAGGCGCGTGAAGAGGATTTATTTGCTCTCCAGATGATTTCTGTAAGTTTGAAGGCAA 4365
Db 442 uThrGlyGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAs 462
QY 4366 TAAAGGTTCTATGTTTGTGATTACAACTTTGCAAGGCGAGCGGAGCATCTTAAC 4425
Db 462 nLysGlySerIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuTh 482
QY 4426 GAAAGAAATTC 4435
Db 482 rLysGluPhe 485
```

## RESULT 39

```
US-10-872-155-325
/ Sequence 325, Application US/10872155
/ Publication No. US20040234536A1
/ GENERAL INFORMATION:
/ APPLICANT: Bhatia, Ajay
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Probst, Peter
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
/ TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
/ FILE REFERENCE: 210121.469C10
/ CURRENT APPLICATION NUMBER: US/10/872,155
/ PRIORITY FILING DATE: 2004-06-18
/ PRIORITY FILING DATE: 09/620,412
/ PRIORITY FILING DATE: 2000-07-20
/ PRIORITY FILING DATE: 09/598,419
/ PRIORITY FILING DATE: 2000-06-20
/ PRIORITY FILING DATE: 09/556,877
/ PRIORITY FILING DATE: 2000-04-19
/ PRIORITY FILING DATE: 09/454,684
/ PRIORITY FILING DATE: 1999-12-03
/ PRIORITY FILING DATE: 09/426,571
/ PRIORITY FILING DATE: 1999-10-22
/ PRIORITY FILING DATE: 09/410,568
/ PRIORITY FILING DATE: 1999-10-01
/ PRIORITY FILING DATE: 09/288,594
/ PRIORITY FILING DATE: 1999-04-08
/ PRIORITY FILING DATE: 09/208,277
/ PRIORITY FILING DATE: 1998-12-08
/ NUMBER OF SEQ ID NOS: 599
/ SOFTWARE: FastSeq for Windows Version 3.0/4.0
/ SEQ ID NO 325
/ LENGTH: 631
/ TYPE: PRT
/ ORGANISM: Chlamydia trachomatis
US-10-872-155-325
```

```
Alignment Scores:
Pred. No.: 3.86e-132 Length: 631
Score: 1590.50 Matches: 316
Percent Similarity: 88.7% Conservative: 7
Best Local Similarity: 86.8% Mismatches: 19
Query Match: 20.2% Indels: 22
DB: 5 Gaps: 4

US-10-701-844-1 (1-4435) x US-10-872-155-325 (1-631)
QY 3391 TTGAGTCGAGCAAGTAGAGTCGGTCTCTAAATAAT-----TGCTTAGATAGTTAA 3441
Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisThrPro----- 146
QY 3442 GTGTTAGCGATGCGCTTTTCTTTGAGATCTACATCATTTTCTTTTATAGCTTGTGTGT 3501
Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCy8 164
QY 3502 TCCTATTTCGTATGATTCGCGAGCTCTCCTCAAGTGTTAACGCTTAATGTAAACGACTCT 3561
Db 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrPro 184
QY 3562 TTTAAGGGGAGACGATGTTTACTTCAATCGAGACTGCGCTTTTGTCAATGTCTTATCAGGA 3621
Db 185 PheLysGlyAspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGly 204
QY 3622 GCTGAAGAAGGTTTCGATTATCTCAGCTAATCGCGACAATTTTAACGATTACCGACAAAAC 3681
Db 205 AlaGluAsnGlySerIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 224
QY 3682 CATCATTTATCATTTACGATTTCTCAAGGGCGAGTTCTTCAAAATTATGCTTCAATTCA 3741
Db 225 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSer 244
QY 3742 CGAGGAGAGACACTTACTCTGAGAGATTTTTCGAGTCTGAGTTCTCGAAAAATGTTTCT 3801
Db 245 AlaGlyGluThrLeuThrLeuLysAspPheSerLeuMetPheSerLysAsnValSer 264
QY 3802 TCGCGAGAAAAAGGAATGATCTCCGGGAAAAACCGTAGTATTTCCGGAGCGAGCGAAGTG 3861
Db 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284
QY 3862 ATTTTCTGGGATACTCCGTCGGGTATTTCTCTTATCTACTGTGCGCAACCTCATCA 3921
Db 285 IlePheThrAspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304
QY 3922 ACTCCGCTGCT-----CCACAGTTAGT 3945
Db 305 ThrProProAlaProAlaProAlaProAlaSerSerSerLeuSerProThrValSer 324
QY 3946 GATGCTCGAAAGGGTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAA 4005
Db 325 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValIys 344
QY 4006 AAAGGGTCATGTCGATAAATCCCGGAATTTCCGAAACAGTTTTCGAGGTAAGAAT 4065
Db 345 LysGlyValMetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSer 364
QY 4066 AATAATAATGCTGGTGGGAGGCGAGTGGGTTCGCTACACCATCAAGTACGACTTTTAC 4125
Db 365 AsnAsnAsnAlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheTh 382
QY 4126 AGTTAAACCTGTAAGGGAAAGTTTCTTTTCACAGATAACGAGTCTTTCGGAGGCGG 4185
Db 382 rValIysAsnCysLysGlyValSerPheThrAspAsnValAlaSerCysGlyGlyG1 402
QY 4186 AGTGGTTTATAAGGCATTGCTTTTCAAGACAAATCAAGGAGGCATATTTTCCGAGG 4245
Db 402 yValValTyrLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgG1 422
QY 4246 GAACACAGCATACGATGATTTAAGGATTTCTGCTCTACTAATCAGGATCAGAATACGGA 4305
```

Db 422 yAsnThrAlaTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrG1 442  
Qy 4306 GACAGAGCGGTGGAGGATTATTGCTCTCCAGATGTTCTGTAAGTTTGAAGGCNA 4365  
Db 442 uThrGlyGlyGlyGlyValIleCysSerProAspAspSerValIlePheGluGlyAs 462  
Qy 4366 TAAAGGTTCTATTGTTTGTATACAACTTTCAAAAGCAGAGCGGAAGCATCCTAAC 4425  
Db 462 nLysGlySerIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuTh 482  
Qy 4426 GAAAGAAATTC 4435  
Db 482 rLysGluPhe 485  
RESULT 40  
US-09-841-260-95  
; Sequence 95, Application US/09841260  
; Publication No. US20030175700A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Probst, Peter  
; APPLICANT: Stromberg, Erika Jean  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS  
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.515  
; CURRENT APPLICATION NUMBER: US/09/841,260  
; CURRENT FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 140  
; SEQ ID NO 95  
; LENGTH: 1016  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis serovar D  
US-09-841-260-95  
Alignment Scores:  
Pred. No.: 6,58e-132 Length: 1016  
Score: 1589.00 Matches: 311  
Percent Similarity: 92.7% Conservative: 5  
Best Local Similarity: 91.2% Mismatches: 10  
Query Match: 20.2% Indels: 15  
DB: 3 Gaps: 2  
US-10-701-844-1 (1-4435) x US-09-841-260-95 (1-1016)  
Qy 3451 ATGCTTTTCTTTGAGACTACATCATTTGTTTGTAGCTGTTGTTGTTCTATTCG 3510  
Db 1 MetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrSer 20  
Qy 3511 TATGGATTCCGAGCTCTCTCAAGTGTAAACGCTTAATGTAACTCCCTCTTTTAAAGGA 3570  
Db 21 TyrGlyPheAlaSerProGlnValLeuThrProAsnValThrThrProPheLysGly 40  
Qy 3571 GACGATGTTTACTTGAATGAGAGCTGCGCTTTTGTCTCAATGCTATGCTATGAGAGCTGAAGAA 3630  
Db 41 AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn 60  
Qy 3631 GGTTCGATTATCTCAGCTAATGCGCAATTTAAACGATTACCGGACAAAACCATACATTA 3690  
Db 61 GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu 80  
Qy 3691 TCATTTACAGATTCTCAAGGCCAGTTCTTCAAAATTTATGCTTCATTTCCAGCAGGAGAG 3750  
Db 81 SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu 100  
Qy 3751 ACATTTACTCTGAGAGATTTTTCAGTCTGATGTTCTCGAAAATGTTTCTTTCGCGAGAA 3810  
Db 101 ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu 120  
Qy 3811 AAGGGAATGATCTCCGGGAAACCGTGAGTATTTCCGAGCAGCGGAAGTATTTCTGG 3870  
Db 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTyr 140  
Qy 3871 GATAACTCCGTGGGGTATTCTCTTTTATCTACTGTGCAACCTCATCATCACTCCGCT 3930

Db 141 AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro 160  
Qy 3931 GCT-----CCAACTAGTTAGTCTGCTCGG 3954  
Db 161 AlaProAlaProAlaProAlaSerSerLeuSerProThrValSerAspAlaArg 180  
Qy 3955 AAAGGCTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAAAAGGGTTC 4014  
Db 181 LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLysGlyVal 200  
Qy 4015 ATGTTTCGATATAATGCTCCGGGAATTTCCGAACAGTTTTTTCGAGGTAAAGATATAATAT 4074  
Db 201 MetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsn 220  
Qy 4075 GCTGTGTGGAGCGAGTGGTTCGCTACACCATCAAGTACGACTTTTACAGTTAAAAA 4134  
Db 221 AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLysAs 238  
Qy 4135 CTGTAAAGGGAAGTTTCTTTTCCAGATAACCTAGCTCTTCGCGAGCGGAGTGTGTTA 4194  
Db 238 nCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyValValty 258  
Qy 4195 TAAAGGCATTGTGCTTTTCAAGAGCAATGAAGAGGCATATTTCTCCGAGGGAACACAGC 4254  
Db 258 rLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAl 278  
Qy 4255 ATACGATGATTTAAGGATTTCTGCTGCTACTAATCAGGATCAGATACGAGACAGAGG 4314  
Db 278 aTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGly 298  
Qy 4315 CGGTGGAGGAGTATTGCTCTCCAGATGATTTCTGTAAAGTTTGAAGGCAATAAAGTTTC 4374  
Db 298 yGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySe 318  
Qy 4375 TATTGTTTGTATTACAACTTTGCAAAAGCGAGCGGAGCATCTTAAACGAAAGATT 4434  
Db 318 rIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPh 338  
Qy 4435 C 4435  
Db 338 e 338

Search completed: May 13, 2006, 12:11:20  
Job time : 896.5 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 13, 2006, 10:31:28 ; Search time 24.6 Seconds  
(without alignments)  
4471.544 Million cell updates/sec

Title: US-10-701-844-1  
Perfect score: 7883  
Sequence: 1 gggcaaaactctcccccg.....gcactcaacgaagaattc 4435

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame+n2p.model -DEV=xlpl  
-Q=/abss/ABSSWEB\_epool/US10701844/runat\_12052006\_165426\_26373/app\_query.fasta\_1  
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs02p -USER=US10701844 @CGN 1 1 71 @runat\_12052006\_165426\_26373  
-NCPV=6 -ICPV=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WRAP TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:.\*  
1: /cgn2\_6/prodata/1/iaa/5 COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/6 COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/H COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/RE COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	66.8	1012	2	US-09-612-402B-2
2	5267	66.8	1012	2	US-09-542-520-2
3	5116	64.9	984	2	US-09-612-402B-43
4	5106.5	64.8	1013	2	US-09-612-402B-15
5	5106.5	64.8	1013	2	US-09-612-402B-16
6	5106.5	64.8	1013	2	US-09-542-520-15
7	5106.5	64.8	1013	2	US-09-542-520-16
8	5090	64.6	1006	2	US-09-556-877-190
9	5090	64.6	1006	2	US-09-620-412C-190
10	5090	64.6	1006	2	US-09-598-419-190
11	5084	64.5	982	2	US-09-556-877-176
12	5084	64.5	982	2	US-09-620-412C-176

13	5084	64.5	982	2	US-09-598-419-176	Sequence 176, App
14	3336.5	42.3	670	2	US-10-197-220-169	Sequence 169, App
15	2552	32.4	505	2	US-09-612-402B-17	Sequence 17, Appl
16	2547	32.3	505	2	US-09-542-520-17	Sequence 17, Appl
17	2350	29.8	458	2	US-09-612-402B-36	Sequence 36, Appl
18	2350	29.8	458	2	US-09-542-520-36	Sequence 36, Appl
19	1735	22.0	325	2	US-09-612-402B-37	Sequence 37, Appl
20	1735	22.0	325	2	US-09-542-520-37	Sequence 37, Appl
21	1590.5	20.2	631	2	US-09-620-412C-325	Sequence 325, App
22	1590.5	20.2	631	2	US-09-598-419-325	Sequence 325, App
23	1589	20.2	1016	2	US-10-197-220-95	Sequence 95, Appl
24	1456	18.5	664	2	US-10-197-220-168	Sequence 168, App
25	1366.5	17.3	999	2	US-09-438-185A-455	Sequence 455, App
26	1359.5	17.2	973	2	US-09-430-723-2	Sequence 2, Appli
27	1129	14.3	928	2	US-09-428-123-2	Sequence 2, Appli
28	1094.5	13.9	967	2	US-09-438-185A-453	Sequence 453, App
29	1092	13.9	949	2	US-09-198-452A-478	Sequence 478, App
30	1083	13.7	930	2	US-09-198-452A-470	Sequence 470, App
31	1079	13.7	938	2	US-09-438-185A-448	Sequence 448, App
32	1041	13.2	947	2	US-09-438-185A-447	Sequence 447, App
33	1021	13.0	937	2	US-09-438-185A-449	Sequence 449, App
34	994.5	12.6	927	2	US-09-198-452A-472	Sequence 472, App
35	991.5	12.6	1414	2	US-09-438-185A-446	Sequence 446, App
36	981.5	12.5	780	2	US-09-438-185A-17	Sequence 17, Appl
37	973	12.3	932	2	US-09-438-185A-6	Sequence 6, Appli
38	964	12.2	922	2	US-09-198-452A-15	Sequence 15, Appl
39	943.5	12.0	866	2	US-09-438-185A-15	Sequence 15, Appl
40	828	10.5	1132	2	US-09-198-452A-466	Sequence 466, App
41	808	10.2	634	2	US-09-438-185A-451	Sequence 451, App
42	804	10.2	643	2	US-09-198-452A-474	Sequence 474, App
43	769.5	9.8	880	2	US-09-556-877-175	Sequence 175, App
44	769.5	9.8	880	2	US-09-620-412C-175	Sequence 175, App
45	769.5	9.8	880	2	US-09-598-419-175	Sequence 175, App

ALIGNMENTS

RESULT 1  
US-09-612-402B-2  
; Sequence 2, Application US/09612402B  
; Patent No. 6642023  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OP INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; Patent No. 6642023  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/09/612,402B  
; CURRENT FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 2  
; TYPE: PRT  
; LENGTH: 1012  
; ORGANISM: Chlamydia sp.  
US-09-612-402B-2

Alignment Scores:  
Pred. No.: 0 Length: 1012  
Score: 5267.00 Matches: 1012  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 66.8% Indels: 0  
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-612-402B-2 (1-1012)

QY 382 ATGCMAAGCTCTTCCATGAGTCTTCTTCAATGATTCCTAGCTTATCTTCTCTCT 441

Db 1 MetGlnThrSerPheHisLysPheLeuSerMetLeuAlaTyrSerCysSer 20

QY 442 TTTAAATGGGGGGATATGAGCAGAAATCATGGTCTCTCAAGGAATTTACGATGGGAG 501  
DB 21 LeuAenGlyGlyTyAlaAlaGluMetValProGlnGlyIleTyAspGlyGlu 40  
QY 502 ACGTAACTGATCATTTCCCTATACCTGTTATAGGAGATCGAGTGGACTACTGTTTTT 561  
DB 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGCAGGAGGTAAACATTAATAAAATCTTGACAAATCTTATGACGCTTTGCCCTTAAGT 621  
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
QY 622 TGTTTTGGGAACCTATTATAGGAGGTTTTACTGTTTTTAGGAGAGACACTCGTTGACTTTC 681  
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGAACATACGACTTCTACAAATGGGGAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741  
DB 101 GluAenIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120  
QY 742 TTTACTATTGAGGCTTTTAAGAAATATCTTTTCCAAATTCATTTACTTATGCGGTA 801  
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
QY 802 CTGCTGCTGCAACGACTTAATAAGGTAGCCAGACTCCGACGACAACTCTACACCGTCT 861  
DB 141 LeuProAlaAlaThrAsnLysGlySerGlnThrProThrThrSerThrProSer 160  
QY 862 AATGGTACTATTATCTAAACAGACTTTTGGTGTACTCAATATGAGAAGTCTCATTC 921  
DB 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnGluLysPheSerPhe 180  
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981  
DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200  
QY 982 GGAAATTACAGCTTTGTGCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys 220  
QY 1042 CAAGTAGTACACAGTTCTCTGCTATAGCTAACGAGGCTCTATTGCTTCTTTGAGCGAAT 1101  
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240  
QY 1102 GTTCAGAGATTAAGGGGGAGGATGCTGCTGTTCAAGATGGGACGAGGAGTGCA 1161  
DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACCTGCGGTAGAGTTT 1221  
DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
QY 1222 GATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAGCTTCTTCTG 1281  
DB 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300  
QY 1282 AATAGTGNAAACCTGTTTCTCAACATGTGCTTCTCTGTTTACATGCTGCTCAAG 1341  
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyIleAlaAlaLys 320  
QY 1342 CAACCAACAAGTGGACAGCTTCTAATACGAGTAATAATTAACGAGATGGAGGACTATC 1401  
DB 321 GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyGlyAspGlyAlaIle 340  
QY 1402 TTCTGTAGAATGGTGGCAACGAGATCCAAATACTCTGGATCAGTTCTCTTCAATGA 1461  
DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly 360  
QY 1462 GAGGAGTAGTTTCTTTAGTAGCAATGTAGTCTGCGGAAAGGGGAGCTATTTATGCC 1521  
DB 361 GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyAla 380  
QY 1522 AAAAAGCTCTCGGTCTAACTGTGGCCCTGTACAAATTTTTAAGGAATATCGCTAATGAT 1581

DB 381 LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp 400  
QY 1582 GGTGGAGCGATTTATTTAGGAGAACTCTGGAGACTCAGTTTATCTGCTGATATGGAGAT 1641  
DB 401 GlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyGlyAsp 420  
QY 1642 ATTATTTTCGATCGGAATCTTAAAGAACAGCAGCAAAGAGAAATGCTGCCGATGTTAATGGC 1701  
DB 421 IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly 440  
QY 1702 GTAACGTGTCTCACAGCCATTTTCGATCGGATCGGAGGAAATAACGACATTAAGA 1761  
DB 441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg 460  
QY 1762 GCTAAAGCAGGCATCAGATTCTCTTAAATGATCCCATCGAGATGGCAACGGAATAAC 1821  
DB 461 AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn 480  
QY 1822 CAGCCAGCGAGTCTTCCAAACTTCTAAATAATTAAACGATCGTGAAGGATACACAGGGAT 1881  
DB 481 GlnProAlaGlnSerSerLysLeuLysIleAsnAspGlyGluGlyTyThrGlyAsp 500  
QY 1882 ATTGTTTTCTTAATGGAACAGTACTTCTTACCAAAATCTTACGATAGACGAGGAAGG 1941  
DB 501 IleValPheAlaAsnGlySerSerThrLeuTyGlnAsnValThrIleGluGlnGlyArg 520  
QY 1942 ATTGTTCTTCTGTAAGGCAAAATATCAGTGAATTTCTTAAGTCACAGAGTGGGAGT 2001  
DB 521 IleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyLysSer 540  
QY 2002 CTGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCACAAACCAACCAACAGCCT 2061  
DB 541 LeuTyMetGluAlaGlySerThrTrpAspPheValThrProGlnProProGlnGlnPro 560  
QY 2062 CCTGCCCTAATCAGTTGATCAGCTTTCCAACTGCAATTTGCTCTTCTTCTTTGTTA 2121  
DB 561 ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeuLeu 580  
QY 2122 GCAACAATGAGTTACGAATCTCTACCAATCTCCAGCGCAAGATTTCTCATCTGCA 2181  
DB 581 AlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisProAla 600  
QY 2182 GTCAATGTTAGCACAACTGCTGTTCTGTTACAATTAAGTGGGCTATCTTTTTTGAGAT 2241  
DB 601 ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAsp 620  
QY 2242 TTGGATGATCAGCTTATGATAGTATGATGCTTGGCTTCTTAATCAAAAATCAATCTC 2301  
DB 621 LeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsnVal 640  
QY 2302 CTGAAATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTTTGACTCTAGG 2361  
DB 641 LeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAspLeuThrLeuGly 660  
QY 2362 AATGAGATGCTAAGTATGGCTATCAAGAAAGCTGGAAGCTTGGCTGGGATCCTTAATCA 2421  
DB 661 AsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsnThr 680  
QY 2422 GCAATAATGTCCTTATCTCTGAAAGCTACATGACTTAAACCTGGAATTAATCTCTGG 2481  
DB 681 AlaAsnAsnGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAsnProGly 700  
QY 2482 CTGAGCAGTAGCTTCTTGTCTCAAAATAGTTTATGGGATCCATCTTTAGATATACA 2541  
DB 701 ProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIleArg 720  
QY 2542 TCTGCGCAATCAGCAATTCAGCAAGTGTGAATGGGCGCTCTTATTTGTCAGGATATGCG 2601  
DB 721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeuTrp 740  
QY 2602 GTTTCGAGGTTTCGAAATTTCTTCTATCATGACCGGATGCTTTTAGTTCAGGATATCGG 2661

741 ValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArg 760  
2662 TATATTAGTGGGGTATTCTCTTAGAGCAAACTCTCTACTTTGGATCATCGATGTTGCT 721  
761 TyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPheGly 780  
2722 CTAGCATTTACCGAAGTATTGCTAGATCTAAAGATTATGTAGTGTGCTTCCCAATCAT 2781  
781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsnHis 800  
2782 CATGCTTGATAGGATCCGTTTATCTATCTACCCAAAGCTTTATGTGATCCTATTG 2841  
801 HisAlaCysIleGlySerValTyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeu 820  
2842 TTCGGAGATCGCTTATCCGTCTAGCTAGCTAGCGGTTTGGGAATCAGCATATGAAACCTCA 2901  
821 PheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSer 840  
2902 TATACATTTGCAGAGAGCGATGTTCTGTGGGATAATACTGTCTGGCTGGAGAGATT 2961  
841 TyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuAlaGlyGluIle 860  
2962 GGAGCGGGATTACCGATTGTGATTACTCCATCTAAGCTCTTATTTGAATGAGTTGCTCCT 3021  
861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuA-gPro 880  
3022 TTCGTCGAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGATCAA 3081  
881 PheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAspGln 900  
3082 GCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCAGTCTCTGTGGAGTGAAGTTT 3141  
901 AlaArgAlaPheLysSerGlyHisLeuLeuAsnLeuSerValProValGlyValLysPhe 920  
3142 GATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGGCGGCTTATCTGTGAT 3201  
921 AspArgCysSerSerThrHisProAsnLysTyrSerPheMetAlaLysTyrIleCysAsp 940  
3202 GCTTATCGCAACCATCTCTGTGATCTGAGACAACTCTCTATCCATCAGAGACATGGACA 3261  
941 AlaTyrArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThr 960  
3262 ACAGATCGCTTTCATTTAGCAACGATGGAGTTGTGTTAGAGATCTATGATGCTTCT 3321  
961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyrAlaSer 980  
3322 CTAAACAGTAATATAGAGTATATGGCCATGGAGATATGATATCGATATCGATGCTTCTCGA 3381  
981 LeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArg 1000  
3382 GGCTATGGTTGATGCGAGGAGTATGAGTCCGGTTC 3417  
1001 GlyTyrGlyLeuSerAlaGlySerArgValArgPhe 1012

## RESULT 2

US-09-542-520-2  
; Sequence 2, Application US/09542520  
; Patent No. 6887843  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John L.  
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 7969-076-999  
; CURRENT APPLICATION NUMBER: US/09/542,520  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/US98/20737  
; PRIOR FILING DATE: 1998-10-01  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1012  
; TYPE: PRT  
; ORGANISM: Chlamydia

## US-09-542-520-2

Alignment Scores:  
Pred. No.: 0 Length: 1012  
Score: 5267.00 Matches: 1012  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 66.8% Indels: 0  
DB: 2 Gaps: 0

## US-10-701-844-1 (1-4435) x US-09-542-520-2 (1-1012)

382 ATGCAAAAGCTCTTTCCATAAGTCTTTCTTTCAATGATCTAGCTTATTTCTGCTCTCT 441  
1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20  
442 TTAATATGGGGGGATATGCACAGAAATCATGTTCTCTCAAGGAATTTACGATGGGGAG 501  
21 LeuAsnGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40  
502 AGCTTAACCTGATCATTTCCCTATCTGTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
562 TCTGCAGAGAGATTAACTATAAAATCTTGACAATTTCTATTCACGCTTTCGCTTTAAGT 621  
61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80  
622 TGTGTTGGAACTTATAGGGAGTTTTTACTGTTTTTAGGAGGAGGACACTCGTTGACTTTC 681  
81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
682 GAGACATACGAGCTTCTACAATGGGCGAGCTCTAAGTAATAGCCCTGCTGATGACTG 741  
101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120  
742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTTCCAATTTGCAATTCATTACTGCGGTA 801  
121 PheThrIleGlyGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140  
802 CTGCTCTGTCGAACGACTAATAGGGTAGCCAGACTCCGACGACACATCTACCCGCTCT 861  
141 LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrThrSerThrProSer 160  
862 AATGTTACTATTATTCTAAACAGATCTTTTCTTACTCAATAATCAGAAAGTTCTCATTC 921  
161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnGlnLysPheSerPhe 180  
922 TATAGTAATTTAGTCTCTCGAGATGGGGAGCTTATAGATGCTTAAGAGCTTAAACGGTTCAA 981  
181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
982 GGAATTAGCAAGCTTTGTGTTTCCAAAGAAATPACTGCTCAAGCTGATGGGGAGCTTGT 1041  
201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
1042 CAAGTAGTCACGATTTCTCTGCTATGGCTTAAAGAGCTCTTATGCTTTGTAGCGAAT 1101  
221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240  
1102 GTTGCAGGAGTAAGAGGGGGAGGATTTGCTGTGTTTCAGGATGGGACGAGGAGTGTCA 1161  
241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnValSer 260  
1162 TCATCTACTTCAACAGAAGATCCAGTAGTAGTTTTCAGAAATATCTCGGTTAGAGTTT 1221  
261 SerSerThrThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
1222 GATGGGAACTAGCCCGGAGTAGGAGGGGATTTACTCTACCGGAAGCTTGTCTTCTCTG 1281  
281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
1282 AATAATGGAAAAACCTTGTGTTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAG 1341

Db 301 AenAenGlyLeuThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaAlaLys 320  
Qy 1342 CAACCAACAGTGGACAGCTTCTTAATACGAGTAAATATACGGAGATGGAGGACTATC 1401  
Db 321 GlnProThrSerGlyGlnAlaSerAenThrSerAenAenTyrGlyAspGlyGlyAlaIle 340  
Qy 1402 TTCTGTAAAGATGGTCCCAAGCAGGATCCAAATACTCTGGATCAGTTTCCTTTGATGGA 1461  
Db 341 PheCysLeuAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAspGly 360  
Qy 1462 GAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGTGGGAAAGGGAGCTATTTATGCC 1521  
Db 361 GluGlyValAlaPhePheSerSerAenValAlaAlaGlyLysGlyAlaIleTyrAla 380  
Qy 1522 AAAAGCTCTCGTTGTCTTAACGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAATGAT 1581  
Db 381 LysLysLeuSerValAlaAenCysGlyProValGlnPheLeuAenAenIleAlaAenAsp 400  
Qy 1582 GGTGAGCGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTTATGGAGAT 1641  
Db 401 GlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGlyAsp 420  
Qy 1642 ATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAATGSC 1701  
Db 421 IleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAenGly 440  
Qy 1702 GTAACTGTCTCTCAAGACCATTTTCGATGGGATCGGAGGAGAAATTAACACATTAAGA 1761  
Db 441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg 460  
Qy 1762 GCTAAGCAGGCGATCAGATTTCTTTAATGATCCCATCGAGATGGCAAAACGGAATTAAC 1821  
Db 461 AlaLysAlaGlyHisGlnIleLeuPheAenAenProIleGluMetAlaAenGlyAenAen 480  
Qy 1822 CAGCAGCGCGATCTTCCAACTCTTAAATTAACGATGGTGAAGGATACACAGCGGAT 1881  
Db 481 GlnProAlaGlnSerSerLysLeuLeuLysIleAenAenGlyGluGlyTyrThrGlyAsp 500  
Qy 1882 ATTGTTTGTGTAAGCAGTACTTTGTACCAAAATGTTTACGATAGACAGGAAGG 1941  
Db 501 IleValPheAlaAenGlySerSerThrLeuTyrGlnAenValThrIleGluGlnArg 520  
Qy 1942 ATTGTTCTTCGAAAGCAAAATATCAGTGAATTTCTTAAGTCAGACAGTGGGAGT 2001  
Db 521 IleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGlySer 540  
Qy 2002 CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACACCACCAAGCCT 2061  
Db 541 LeuTyrMetGluAlaGlySerThrTrpAspPheValThrProGlnProGlnGlnPro 560  
Qy 2062 CCTGCGCTTAATCAGTTGATCAGCTTTTCCAACTCGCATTTGTCTTCTTCTTTGTTA 2121  
Db 561 ProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeuLeu 580  
Qy 2122 GCAACAAATGAGTACGAATCTCTCAATCTCCAGCGAAGATTTCTCATCTGCA 2181  
Db 581 AlaAenAenAlaValThrAenProProThrAenProProAlaGlnAenSerHisProAla 600  
Qy 2182 GTCATTTGGTAGCACACTGCTGTTCTGTGTAAATTAGTGGCCCTATCTTTTTCGAGT 2241  
Db 601 ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAsp 620  
Qy 2242 TTGGATGATACAGCTATGATAGGTATGATGGCTAGTTCCTAATTAGTGGCCCTATCTTTTCGAGT 2301  
Db 621 LeuAenAenThrAlaTyrAspArgTyrAspTyrLeuGlySerAenGlnLysIleAenVal 640  
Qy 2302 CTGAAATTAACGTTAGGAGTAAAGCCCCAGCTAATAGCCCCATCAGATTTGATCTAGGG 2361  
Db 641 LeuLysLeuGlnLeuGlyThrLysProProAlaAenAlaProSerAspLeuThrLeuGly 660  
Qy 2362 AATGAGATGCTTAAGTATGCTATCAAGGAAGCTGGAAGCTGGGAGTCTTAATACA 2421

Db 661 AenGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTyrAspProAenThr 680  
Qy 2422 GCAAAATATGCTCTTATATCTCTGAAAGCTACATGGACTAAACCTGGTATAATCTCTGG 2481  
Db 681 AlaAenAenGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnProGly 700  
Qy 2482 CCTCAGCGAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATACGA 2541  
Db 701 ProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIleArg 720  
Qy 2542 TCTCGCATTTCAAGCAATTTCAAGCAAGTGTGGATGGCGCTCTTATTTGCGAGGATATGG 2601  
Db 721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeuTrp 740  
Qy 2602 GTTTCGAGTTTCGAATTTCTTCTATCATGACCGCGATGCTTTAGTCAAGGATATCGG 2661  
Db 741 ValSerGlyValSerAenPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArg 760  
Qy 2662 TATATTAGTGGGCTTATCTCTTAGGAGCAAACTCTTCTTGGATCATCATGTTTGGT 2721  
Db 761 TyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPheGly 780  
Qy 2722 CTAGCATTTACCGAAGTATTTGGTAGATCTTAAAGATATATGATGTGTCTGTTCCAATCAT 2781  
Db 781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAenHis 800  
Qy 2782 CATGCTTCATAGATCGTTTATCTATCTACCAACAAAGCTTTATGTGGATCTCTATTTG 2841  
Db 801 HisAlaCysIleGlySerValTyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeu 820  
Qy 2842 TTCGAGATGGTTCCTCCGCTAGCTACCGGTTTGGGAATCAGCATATCAAAACCTCA 2901  
Db 821 PheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAenGlnHisMetCysThrSer 840  
Qy 2902 TATCATTTGCAAGGAGGAGCGATTTCTGTTGGGATATAAATCTGCTCGCTGGAGAGAT 2961  
Db 841 TyrThrPheAlaGluGluSerAspValArgTrpAenAenCysLeuAlaGlyGluIle 860  
Qy 2962 GGAGCGGATTTACCGATTGTGATTTACTCCATCTAAGCTCTATTTGAATGAGTTGGCTCT 3021  
Db 861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArgPro 880  
Qy 3022 TTGCTGCAAGCTGAGTTTCTTATCCGATCATGATCTTTTACAGAGAGGCGCATCAA 3081  
Db 881 PheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAspGln 900  
Qy 3082 GCTCGGCGATTTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTT 3141  
Db 901 AlaArgAlaPheLysSerGlyHisLeuLeuAenLeuSerValProValGlyValLysPhe 920  
Qy 3142 GATCATGTTCTTAGTACATCATCTTAATAATATAGCTTTTATGGCGGCTTATATCTGTAT 3201  
Db 921 AspArgCysSerSerThrHisProAenLysTyrSerPheMetAlaAlaTyrIleCysAsp 940  
Qy 3202 GCTTATCCGACCATCTCTGGTACTGAGACACGCTCTATCCATCAAGAGACATGGACA 3261  
Db 941 AlaTyrArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThr 960  
Qy 3262 ACAGATGCTTTCATTTAGCAAGCATCGAGTTGTGGTTAGAGGATCTATGTATGCTTCT 3321  
Db 961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyrAlaSer 980  
Qy 3322 CTAACAAATATATAGATATATGGCCATGGAAAGATATGATATCGAGATGCTTCTGCA 3381  
Db 981 LeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArg 1000  
Qy 3382 GCCTATGCTTTTCAGTGCAGGAGTAGCTCCGGTTC 3417  
Db 1001 GlyTyrGlyLeuSerAlaGlySerArgValArgPhe 1012

RESULT 3

US-09-612-402B-43

; Sequence 43, Application US/09612402B

```
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-43

Alignment Scores:
Pred. No.: 0 Length: 984
Score: 5116.00 Matches: 984
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.9% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-612-402B-43 (1-984)

QY 466 GAAATCATGTTCTCTCAAGGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTAT 525
DB 1 GluilewValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20

QY 526 ACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTTCTGCGAGAGATTAAATTA 595
DB 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40

QY 586 AATCTTGACAAATCTATTGCGCTTTGCCCTTAAAGTTGTTTGGGAACTATTAGGGAGT 645
DB 41 AsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60

QY 646 TTTACTGTTTTAGGAGAGACACTCGTTGACTTTTCGAGAACATACGACTTCTACAAAT 705
DB 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80

QY 706 GGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGCTTTTAAAGAA 765
DB 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100

QY 766 TTATCTTTTCCAAATTCATTTACTTTCGCGTACTGCTGCTGCAACGACTAATAAG 825
DB 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120

QY 826 GSTAGCCAGACTCCGAGACCAACATCTACACCGTCTAATGCTACTATTTTCTTAAACA 885
DB 121 GlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTyrSerLysThr 140

QY 886 GATCTTTTGTACTCAATAATGAGAAGTTCTCAATTTCTATAGTAATTTAGTCTCTGGAGAT 945
DB 141 AspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAsp 160

QY 946 GGGGGAGCTATAGATGCTTAAGACTTAACGGTTCAAGGAAATAGCAAGCTTTGTCTTC 1005
DB 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180

QY 1006 CAAGAAATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCAACAGTTTCTCTGCT 1065
DB 181 GlnGluAsnThrAlaGlnAlaAspGlyCysGlnValValThrSerPheSerAla 200

QY 1066 ATGGCTTAACGAGGCTCTATTGCTCTTTGAGCGAATGTTGAGGAGTAAGGGGGAGGG 1125
DB 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
```

```
QY 1126 ATTGCTGCTGTTTCCAGATGGGAGGAGGTGTCTCATCTTCAACAGAGATCCA 1185
DB 221 IleAlaAlaValGlnAspGlyGlnGlyValSerSerSerThrSerThrGluAspPro 240

QY 1186 GTAGTAAGTTTTCCAGAAATCTCGGTAGAGTTTGTATGGGAACTAGCCGAGTAGGA 1245
DB 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260

QY 1246 GGAGGGATTTACTCTCTACGGGAACTGCTTCTTGAATAATATGGAATAACCTTGTTC 1305
DB 261 GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280

QY 1306 AACAATGTTGCTTCTCTCTTACATTTGCTGTCTAAGCAACCAAGTCGAGCGTTCT 1365
DB 281 AsnAsnValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300

QY 1366 AATACGAGTAATAATTACGGAGATGGAGAGCTATCTTCTGTAAAGATCGTCGCAAGCA 1425
DB 301 AsnThrSerAsnAsnTyrGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320

QY 1426 GGATCCAATAACTCTGGATCAGTTTCTTGTATGGAGGGAGTAGTATTTTCTTAGTAGC 1485
DB 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340

QY 1486 AATGTAGCTGCTGGGAAAGGGGAGCTATTTATGCCAAAAGCTCTCGGTGCTAACTGT 1545
DB 341 AsnValAlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCys 360

QY 1546 GGCCCTGTACAATTTTAAAGGAATATCGCTAATATGATGGTGGAGCGATTTATTAGAGAA 1605
DB 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380

QY 1606 TCTGGAGAGCTCAGTTTATCTGCTGATTATGAGATATTATTTTCGATGGGAATCTTAAA 1665
DB 381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400

QY 1666 AGAACACCAAGAGAGATCTGCCGATGTTAATGGCGTAACTGTGTCTCACAAAGCAAT 1725
DB 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420

QY 1726 TCGATGGGATCGGGAGGGAATAATACGACATTAAGAGCTTAACGAGCGCATCAGATCTC 1785
DB 421 SerMetGlySerGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440

QY 1786 TTTAATGATCCCATCCAGATGCCAACGGAAATAACACGACGCGAGCTCTTCCAAACTT 1845
DB 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerLysLeu 460

QY 1846 CTAATAATTAACGATCGTGAAGGATACACAGGGGATATTGTTTTTCTAATGGAAGCAGT 1905
DB 461 LeuLysIleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480

QY 1906 ACTTTGTACCAAAATGTTACGATAGACGAGGAGGATTTGTTCTTGTGAAAGGCAAAA 1965
DB 481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500

QY 1966 TTATCAGTGAATCTCTAAGTACAGACGAGTGGGAGTCTGTATATGGAAGCTGGAGTAGCA 2025
DB 501 LeuSerValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThr 520

QY 2026 TGGGATTTTGTAATCTCACAAACCAACCAACGCTCTCCGCTAATCAGTTGATCAGC 2085
DB 521 TrpAspPheValThrProGlnProProGlnGlnProProAlaAlaAsnGlnLeuIleThr 540

QY 2086 CTTTCCAATCTGATTTGCTCTTCTTCTTGTGTAGCAAAACAATGAGTTACGAATCCT 2145
DB 541 LeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnPro 560

QY 2146 CCTACCAATCTCCAGCGCAAGATTTCTCATCTCGAGTCATCTGTAGCACAACTGCTGGT 2205
DB 561 ProThrAsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGly 580

QY 2206 TCTGTTACAATTAGTGGGCCTATCTTTTGGAGGATTTGGATGATACAGCTTATGATAGG 2265
```

```
Db 501 SerValThrIleSerGlyProIlePheGluAspLeuAspPheAlaTyrAspArg 600
QY 2266 TATGATGGCTAGGTTCTTAATCAAAAAATCAATGCTCTGAAATTAACAGTAGGGACTAAG 2325
Db 601 TyrAspTyrLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLys 620
QY 2326 CCCGACGCTAATGCCCATCAGATTGACTCTAGGGAATGAGATCCCTAAGTATGGCTAT 2385
Db 621 ProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyr 640
QY 2386 CAAGAACTGGAAGCTTCGCGGATCCTTAATACAGCAAAATAATGTCCTTATATCTGTG 2445
Db 641 GlnGlySerTyrPheLeuAlaTyrAspProAsnThrAlaAsnAsnGlyProTyrThrLeu 660
QY 2446 AAAGCTACATGACTGACAAAACCTGGGTATATCTCGGCTCGAGCGAGTAGCTTCTTGGTT 2505
Db 661 LysAlaThrTyrThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuVal 680
QY 2506 CAAATAGTTATGGGATCCATTTTAGATATATACGATCTGCCATTCAGCAATCAAGCA 2565
Db 681 ProAsnSerLeuTyrPheGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAla 700
QY 2566 AGTGTGGATGGCGCTCTTATTGTCGAGGATTATGGTTCTCGAGATTTCGAATTTCTTC 2625
Db 701 SerValAspGlyArgSerTyrCysArgGlyLeuTyrPheValSerGlyValSerAsnPhePhe 720
QY 2626 TATCATGACCGGATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTTATTCCTTA 2685
Db 721 TyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeu 740
QY 2686 GGAGCAAACTCTACTTTGGATCATCGATGTTGGTCTAGCATTTACCGAAGTATTTGGT 2745
Db 741 GlyAlaAsnSerTyrPheGlySerMetPheGlyLeuAlaPheThrGluValPheGly 760
QY 2746 AGATCTAAGATTATGATGTCGTCTCAATCATCATCGTTGTCATAGGATCCGTTTAT 2805
Db 761 ArgSerLysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyr 780
QY 2806 CTATCTACCAACAAGCTTTATGGATCTTATTTGTCGGAGATCGGTTATCCGTGT 2865
Db 781 LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAla 800
QY 2866 AGCTACGGGTTGGGAATCAGCATATGAAACCTCATATACATTTTCGAGAGAGCGAT 2925
Db 801 SerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluSerAsp 820
QY 2926 GTTCGTTGGGATAATAACTGTCGTGGTGAGAGATTGGAGCGGATTACCGATTGTGATT 2985
Db 821 ValArgTyrAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840
QY 2986 ACTCATCTAAGCTCTATTGATGAGTTGCGTCCTTTTCGTCGAAGTGAGTTTCTTAT 3045
Db 841 ThrProSerLysLeuTyrLeuAsnGlnLeuArgProPheValGlnAlaGluPheSerTyr 860
QY 3046 GCCGATCATGAATCTTTTACAGAGAGCGGATCAAGCTCGGGCATTTCAAGACGGCAT 3105
Db 861 AlaAspHisGluSerPheThrGluGlyAspGlnAlaArgAlaPheLysSerGlyHis 880
QY 3106 CTCCTAAATCATAGTTCCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCT 3165
Db 881 LeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisPro 900
QY 3166 AATAAATATAGCTTTATGCGGCTTATATCTGTGATGCTTATCGCACCACTCTCGTACT 3225
Db 901 AsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThr 920
QY 3226 GAGACAACTCTTATCCCATCAAGACATGGAACAGATGCCCTTCATTTAGCAGA 3285
Db 921 GluThrThrLeuLeuSerHisGlnGluThrTyrThrAspAlaPheHisLeuAlaArg 940
QY 3286 CATGGATTGTGGTATAGGATCTATGATGCTTCTCTAACAGTAATATAGAGTATAT 3345
```

```
Db 941 HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyr 960
QY 3346 GGCCATGGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTCAGTCAGGAAGT 3405
Db 961 GlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980
QY 3406 AGAGTCCGGTTC 3417
Db 981 ArgValArgPhe 984
RESULT 4
US-09-612-402B-15
; Sequence 15, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-15
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.4% Conservative: 17
Best Local Similarity: 96.7% Mismatches: 15
Query Match: 64.8% Indels: 1
DB: 2 Gaps: 1
US-10-701-844-1 (1-4435) x US-09-612-402B-15 (1-1013)
QY 382 ATGCAAAAGCTCTTCCATAAGTCTTCTTCAATGATTCTAGCTTATTCTTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATGGGGGGGATATGACGAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuAsnGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGTATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGAGTTAAACATTAAATAATCTTCACAAATCTTATTCAGCTTTGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
QY 622 TGTTTTGGGAACCTTATTAGGGAGTTTACTGTGTTTTAGGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTAGGGTTTAAAGAAATTATCTTTTCCAAATTGCAATTCATTACTTGCCTGA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140
QY 802 CTGCGCTGCTCAACCACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACCGCTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrSerThrProSer 160
```





Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900  
QY 3079 CAAGCTCGGCATTCACAGAGCGGACATCTCTAAATCTATCATGTTCCCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyVallys 920  
QY 3139 TTGATCGATGTTCTAGTACACATCCTAATAATATAGCTTTATGGCGGCTTATCTGTT 3198  
Db 921 PheAspArgCysSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940  
QY 3199 GATGCTTATCGCACCATCTCTGGTACTCGACACAGCTCTCCATCCATCAAGAGACATGG 3258  
Db 941 AspalatyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTTTTCATTAGCAAGACATGGAGTTGTGGTTAGGAGTCTATGATGCT 3318  
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980  
QY 3319 TCTCTAACAGTAATATATAGATATATGCGCATCGAAGATATGATATCGAGATGCTTCT 3378  
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000  
QY 3379 CGAGCTATGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417  
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

## RESULT 5

US-09-612-402B-16  
; Sequence 16, Application US/09612402B  
; Patent No. 6642023  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; Patent No. 6642023  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/09/612,402B  
; CURRENT FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-09-612-402B-16

Alignment Scores:  
Pred. No.: 0 Length: 1013  
Score: 5106.50 Matches: 980  
Percent Similarity: 98.2% Conservative: 15  
Best Local Similarity: 96.7% Mismatches: 17  
Query Match: 64.8% Indels: 1  
Gaps: 2

US-10-701-844-1 (1-4435) x US-09-612-402B-16 (1-1013)

QY 382 ATGCAACGCTCTTCCATAAGTCTCTTCTCAATGATTTCTAGCTATTTCTTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20  
QY 442 TTAATGGGGGGGATATGCACAGAAATCATGTTCTCTCAAGGAATTTACATGGGAG 501  
Db 21 LeuThrGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40  
QY 502 ACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGCGAGGAGGTAAACATTAATAAATCTTGACAATCTTATTCGAGCTTTGCTTTAAGT 621

Db 61 SerAlaGlyGluLeuThrLeuLysAenLeuAspAenSerIleAlaAlaLeuProLeuSer 80  
QY 622 TGTGTTGGGAACCTTATTAGGAGGTTTACTGTGTTTTAGGAGAGACACTCTGTTGACTTTC 681  
Db 81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGAACATACGAGCTTCTACAAATGGGGAGCTCTAAGTAATAGCGCTGCTGATGGAGCTG 741  
Db 101 GluAenIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120  
QY 742 TTTACTATGAGGCTTTTAAAGAAATATCTCTTTTCCAAATTCGAATTCATTACTTCCCGTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140  
QY 802 CTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACAACTCTACACCGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160  
QY 862 AATGGTACTATTTATCTTAAACAGATCTTTTGTGTACTCAATAATGAGAAGTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAenAsnGluLysPheSerPhe 180  
QY 922 TATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAAGAGCTTAAACGGTTCAA 981  
Db 181 TyrSerAenLeuValSerGlyAspGlyGlyThrIleAspAlaLysSerLeuThrValGln 200  
QY 982 GGAATTAGCAAGCTTTGTGCTTCCAAAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAenThrAlaGlnAlaAenGlyGlyAlaCys 220  
QY 1042 CAAGTGTACCACTTTCTGCTATGCTAAACAGGCTCTATTGCTTGTAGCGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240  
QY 1102 GTTCAGGAGTAAGAGGGAGGAGGATGCTCTCTGTTCAGGATGGCGACAGGAGTGTCA 1161  
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAenGlyGlnGlyValSer 260  
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATATCTGCGGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280  
QY 1222 GATGGACGTAGCCCGAGTAGGAGGAGGATTTACTCTCTACGGGAAAGTGTCTTCTCTG 1281  
Db 281 AspGlyAenValAlaArgValGlyGlyIleTyrSerTyrGlyAenValAlaPheLeu 300  
QY 1282 AATAATGGAACCTTGTCTTCTCAAAATGTTGCTCTCTCTGTTTACATTGCTGCTAAG 1341  
Db 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaAlaGlu 320  
QY 1342 CAACCAACAGTGCACAGGCTTTCTAATACGAGTAATAATTACGGAGATGGAGGAGCTATC 1401  
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAspAenTyrGlyAspGlyGlyAlaIle 340  
QY 1402 TTCTGTAGAAATGTCGGCAA--GCAGATCCCAATACCTCGATCAGTTTCTCTTCAT 1458  
Db 341 PheCysLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAsp 360  
QY 1459 GGAGAGGAGTAGTTTTCTTTAGTAGCAATGTAGCTGTGGGAAAGGGGAGCTATTTAT 1518  
Db 361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTyr 380  
QY 1519 GCCAAAAAGCTCTCGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGAAATATCCCTAAT 1578  
Db 381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400  
QY 1579 GATGTTGAGCGAATTTATTTAGGAATCTCGAGAGCTCAGTTTATCTGCTGATGATGGA 1638  
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
QY 1639 GATATTTTTCGATGGGAATCTTAAAGAACAGCAAGAGAAATGCTGCCGATGTTAAT 1698  
Db 421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440

QY	1699	GGCTAAGTCTGCTCTCAAGCCATTTCCGATGGGATCGGAGGGAATAACGACATTA	1758	QY	2779	CATCATGCTGCATAGGATCCGTTTATCTATCTATCTACCAACAGCTTTATGGGATCCTAT	2838
DB	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrLeu	460	DB	801	HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr	820
QY	1759	AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT	1818	QY	2839	TTGTTCCGAGATCGGTTTATCCGTCCTAGCTACGGGTTTGGGAATCAGCATATGAAACC	2898
DB	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480	DB	821	LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr	840
QY	1819	AACAGCCAGCCAGCTCTTCCAACTTCTAAAAATTAACGATGGTGAAGGATACACAGG	1878	QY	2899	TCATATACATTTGCAGAGGAGGATGTTTCTGTTGGGATAATACTGTCTGGCTGGAGAG	2958
DB	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly	500	DB	841	SerTyrThrPheAlaGluGluSerAspValArgTProAsnAsnAsnCysLeuValGlyGlu	860
QY	1879	GATATTGTTTCTTAATCGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGCA	1938	QY	2959	ATTGAGCGGATTAACGATTGTGATTACTCCATCTAAGCTCTATTGAATGAGTGGCT	3018
DB	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520	DB	861	IleGlyValGlyLeuProIleValThrProSerLysLeuTyrLeuAsnGluLeuArg	880
QY	1939	AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG	1998	QY	3019	CTTTCTGTCGACGCTGAGTTTCTTATCCGATCATGAATCTTTTACAGAGGAAGCGAT	3078
DB	521	ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540	DB	881	ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp	900
QY	1999	AGTCTGTATATGAAGCTGGAGTACATGGGATTTTGTAACTTCCCAACACACACACAG	2058	QY	3079	CAAGCTCCGGGCATTCAGAGCGGACATCTCTAAATCTATCAGTTCCTCTTGGAGTGAAG	3138
DB	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560	DB	901	GlnAlaAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys	920
QY	2059	CTCTCTGCCCTAATCAGTTGATCAGCTTCTCAATCTGCATTTGCTCTTCTTCTTTG	2118	QY	3139	TTTGATCATGTTTCTAGTACACATCTCTAATAATATAGCTTTATGGCGGCTTATATCTGT	3198
DB	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580	DB	921	PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys	940
QY	2119	TTAGCAAAACAATGACGATACGATCTCTTACCAATCTCCAGCGCAAGATTCATCCT	2178	QY	3199	GATGCTTATCGCACCATCTCTGCTACTGAGCAACGCTCTCTATCCCATCAAGAGACATGG	3258
DB	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600	DB	941	AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp	960
QY	2179	GCAGTCATTGGTAGCAACTGCTGTTCTGTACAAATTAGTGGGCTATCTTTTGTAG	2238	QY	3259	ACAAACAGATGCTCTTCATTAGCAAGACATGCGATTGTGCTTAGAGGATCTATGTATGCT	3318
DB	601	AlaValIleGlySerThrAlaGlyProValThrIleSerGlyProPhePheGlu	620	DB	961	ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla	980
QY	2239	GATTGGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT	2298	QY	3319	TCTTAACAAAGTAATATAGATATATAGCCATATGCGCATATGAGATATGAGATGCTTCT	3378
DB	621	AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp	640	DB	981	SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer	1000
QY	2299	GTCTCAAAATTAAGTATGAGGATTAAGCCCAAGCTAATGCCCACTCAGATTTGACTCTA	2358	QY	3379	CGAGGCTATGTTGATGTCAGAGTAGAGTCCGGTTC	3417
DB	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660	DB	1001	ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe	1013
QY	2359	GGGAATGAGATGCTTAAGTATGCTATCAAGAGCTGGAAGCTGCTGGGATCTCTAAT	2418	RESULT 6			
DB	661	GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn	680	US-09-542-520-15			
QY	2419	ACAGCAATATGCTTATCTCTGAAAGCTACATGGAATAAAGCTGGTATATCTCT	2478	; Sequence 15, Application US/09542520			
DB	681	ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro	700	; Patent No. 6887843			
QY	2479	GGGCTGAGGAGTCTCTTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATA	2538	; GENERAL INFORMATION:			
DB	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle	720	; APPLICANT: Jackson, W. James			
QY	2539	CGATCTGCGCATTCAGCAATTCAGCAAGTGTGATGGGCTCTTATTTGCGAGGATTA	2598	; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF			
DB	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740	; FILE REFERENCE: 7969-076-999			
QY	2599	TGGGTTCTGAGTTCGAATTTCTCTATCATGACCGCATGCTTTAGCTGAGGATAT	2658	; CURRENT APPLICATION NUMBER: US/09/542,520			
DB	741	TrpValSerGlyValSerAsnPheSerTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760	; CURRENT FILING DATE: 2000-04-03			
QY	2659	CGGTATATTAGTGGGGTTATCTCTAGGAGCAACTCTTACTTTGGATCATCGATGTTT	2718	; PRIOR APPLICATION NUMBER: PCT/US98/20737			
DB	761	ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780	; PRIOR FILING DATE: 1998-10-01			
QY	2719	GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGATGATGCTGCTTCAAT	2778	; NUMBER OF SEQ ID NOS: 41			
DB	781	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn	800	; SOFTWARE: Patent In Ver. 2.0			
				; SEQ ID NO 15			
				; LENGTH: 1013			
				; TYPE: PRT			
				; ORGANISM: Chlamydia			
				US-09-542-520-15			
				Alignment Scores:			
				Pred. No.: 0		Length: 1013	
				Score: 5106.50		Matches: 980	
				Percent Similarity: 98.4%		Conservative: 17	
				Best Local Similarity: 96.7%		Mismatches: 15	
				Query Match: 64.8%		Indels: 1	
				DB: 2		Gaps: 1	

US-10-701-844-1 (1-4435) x US-09-542-520-15 (1-1013)

QY 382 ATGCAACGCTCTTCCATAAGTCTCTTCTTTCAATGATCTAGCTTATCTTGCTGCTCT 441  
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20  
QY 442 TTAATGGGGGGGATATGCGACAGAAATCATGGTTCCTCAAGGAATTCAGATGGGAG 501  
Db 21 LeuAsnGlyGlyGlyAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40  
QY 502 ACGTTAACTGATCATTTCCCTATACGTATAGGAGATCCGAGTGGGACTACTGTTTTT 561  
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60  
QY 562 TCTGCGAGAGAGTTAAACATTTAAAAATCTTGCAATTCATTGCGAGCTTTGCTTTAAGT 621  
Db 61 SerAlaGlyGlyLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80  
QY 622 TGTTTTGGGAACCTTATTAGGAGTTTTTACTGTTTTAGGAGAGGACACTCGTTGACTTTC 681  
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100  
QY 682 GAGACATACGACCTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741  
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120  
QY 742 TTACTATTAGGGTTTTAAAGAATTTATCCTTTTCCAAATGCAATTCATTACTTCCCGTA 801  
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140  
QY 802 CTGCTGCTGCACGACTTAATAAGGGTAGCGAGCTCCGACGACCAACATCTACACGGTCT 861  
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrSerThrProSer 160  
QY 862 ATGTGACTATTATTCTAAACAGATCTTTTGTACTCAATAATAGAGAGTTCTCATTC 921  
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180  
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981  
Db 181 TyrSerAsnSerValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200  
QY 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041  
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220  
QY 1042 CAAGTAGTCACCAAGTTCTCTGCTATGGCTAACGAGGCTCTATTGCTTGTACGAAT 1101  
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240  
QY 1102 GTTGCGAGGATGAAGGGGGAGGATGCTGCTGTTCCAGATGGCGCAGCGAGGTGCA 1161  
Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260  
QY 1162 TCATCTACTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACTCGCGTAGAGTTT 1221  
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280  
QY 1222 GATGGGAAGTACCGGAGTAGGAGGATTTACTCTACGAGGCTCTATTGCTTGTCTTCCTG 1281  
Db 281 AspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300  
QY 1282 AATAATGGAAAAACCTTTGTTCTCAACAAATGTTGCTTCTCTGTTTACATGCTGCTAAG 1341  
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320  
QY 1342 CAACCAACAAAGTGACAGCTTCTTAATACAGTAATAATACGAGATCGGAGAGCTATC 1401  
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340  
QY 1402 TTCTGTGAAGATGGTGGCGAA--GCAGGATCCAATAACTCTGGATCAGTTTCCTTTGAT 1458  
Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360

QY 1459 GGAGGAGGAGTAGTTTCTTTTAGTAGCAATGTAAGTCTGCGAAAGGGGAGCTATTAT 1518  
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380  
QY 1519 GCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTTAAGCAATATCCCTAAT 1578  
Db 381 AlalysLysLeuSerValAlaAsnCysGlyProValGlnLeuLeuGlyAsnIleAlaAsn 400  
QY 1579 GATGTCGAGCGCATTTATTAGGAGAACTCTCGAGAGCTCAGTTTATCTGCTGATTATGGA 1638  
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420  
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAAACAGCAAGAGAGATGCTGCGGATGTTAAT 1698  
Db 421 AspMetIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440  
QY 1699 GCGCTAACTGTCCTCACAGCCATTTTCGATCGGATCGGAGGGGAAATAACGACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460  
QY 1759 AGAGCTAAAGCAGCGCATCAGATTCTCTTTAATGATCCATCGAGATGGCAACGAAAT 1818  
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480  
QY 1819 AACCGCAGCGCAGTCTTCCAAATCTTCAAAATTAACGATGTTGAAGGATACACAGG 1878  
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500  
QY 1879 GATATTGTTTTGCTTAATGGAAGCAGTACTTTGTACCAAAATGTTACCATAGACAGAGG 1938  
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGlnGlnGly 520  
QY 1939 AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGTGGG 1998  
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540  
QY 1999 AGTCGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCCAACACCACCAACAG 2058  
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560  
QY 2059 CCTCGTCGGCTTAATCAGTTGATCAGCTTCCCATCTGCAATGCTGCTCTCTTCTCTTTG 2118  
Db 561 ProProAlaAlaAsnGlnSerIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580  
QY 2119 TTAGCAAAACAATGCAAGTTACGAATCTCTTACCATCTCCAGCGCAAGATTTCTCATCT 2178  
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600  
QY 2179 CGAGTCATGCTAGCAACCTGCTGGTCTGTATACAAATTAGTGGGCTATCTTTTGTAG 2238  
Db 601 AlaValIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620  
QY 2239 GATTTGGATGATACAGCTTATGATAGGTATGATGCTAGTTCGTTCTTAATCAAAAAATCAAT 2298  
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640  
QY 2299 GTCTCGAAATACAGTTAGGGACTTAAGCCCCCAGCTTAATGCCCCATCAGATTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAsnAlaProSerAspLeuThrLeu 660  
QY 2359 GGAATAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTGCGTGGGATCCTAAT 2418  
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680  
QY 2419 ACAGCAAAATATGTCCTTACTCTGAAGCTACATGAGCTTAAAGCTGGGTATATCCT 2478  
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrThrThrLysThrGlyTyrAsnPro 700  
QY 2479 GGGCCTGAGCGAGTAGTCTTCTTGGTTCCAATAGTTTATGGGATGCCATTTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720

```
QY 2539 CGATCTGCGCATTCAGCAATTCACCAAGTGTGGATGGCGCTCTTATTGTGAGGATTA 2598
Db |||
QY 721 ArgSerAlaHisSerAlaGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Db |||
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTTATCATGACCGCGATGCTTTAGGTGAGGATAT 2658
Db |||
QY 741 TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCCTACTTTGGATCATCGATGTTT 2718
Db |||
QY 761 ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCAATTTACCGAAGTATTTGTAGATCTAAAGATTATGTAGTGTGCTTCCCAAT 2778
Db |||
QY 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTGATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTGGATCCTAT 2838
Db |||
QY 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTCCGAGATCGCTTTATCCGTCTAGCTACGAGTTCGGGTAATCAGCATATGAAACC 2898
Db |||
QY 821 ValPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGGAGCGATGTTCTGTTGGGATAATAACTGTCTGGCTGGAGAG 2958
Db |||
QY 841 SerTyrThrPheAlaGluGluSerAspValCysTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGGAGCGGGATTACCGGATTGTGATTCTTCAATCTAAGCTCTTATTGAATGAGTTGCGT 3018
Db |||
QY 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CTTTCGTCGACAGCTGAGTTTCTTATGCGGATCATGAATCTTATTGAATGAGTTGCGT 3078
Db |||
QY 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAGCTCGGCATTCAGAGCGGACATCTCTAAATCTATCAGTCTCTGTTGGAGTGAAG 3138
Db |||
QY 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGGCGCTTATCTGT 3198
Db |||
QY 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATGCAACCATCTCTGTACTGAGACAAACGCTCTCTATCCCATCAGAGCATGG 3258
Db |||
QY 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATCCCTTCAATTTAGCAACATGGAGTGTGCTGTAGAGATCTATGTATGCT 3318
Db |||
QY 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAAATAGAGTATATGGCCATCGGAGATGATGAGTATCGATGCTTCT 3378
Db |||
QY 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGTTGATGTCAGGAGTATGAGTCCGGTTC 3417
Db |||
QY 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

RESULT 7

```
US-09-542-520-16
; Sequence 16, Application US/09542520
; Patent No. 6887843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; PRIOR FILING DATE: 1998-10-01
```

```
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-542-520-16

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.2% Conservative: 15
Best Local Similarity: 96.7% Mismatches: 17
Query Match: 64.8% Indels: 1
DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x US-09-542-520-16 (1-1013)
QY 382 ATGCAACGCTCTTTCCATAAGTCTTTCTTCTTCAATGATCTAGCTTATTCTTGCTGCTCT 441
Db |||
QY 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATATGGGGGATATGCACAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501
Db |||
QY 21 LeuThrGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGTATCATCTTCCCTATCTACTGTATAGGAGATCCGAGTGGGACTACTGTTTTT 561
Db |||
QY 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCGAGGAGTAAACATTAATAAATCTTGACAAATTTGCAATTTATTCAGCTTTGCTTTAAGT 621
Db |||
QY 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTTTGGAACTTATAGGGAGTTTTACTGTTTTAGGAGGAGGACACTCGTTGACTTTC 681
Db |||
QY 81 CysPheGlyAsnLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGACATACGCACTTCTACAAATGGCGAGCTCTAAGTAATAGACGCTGTGATGAGACTG 741
Db |||
QY 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTTCCAATTTGCAATTTCAATTTGCGGTA 801
Db |||
QY 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGATTAATAGGTAGCCAGACTCCGACGACACATCTACACGCTCT 861
Db |||
QY 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTCTAAACAGATCTTTTGTACTCAATAATGAGAAAGTTCTCATTTTC 921
Db |||
QY 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGAGATGGGGAGCTTATAGATGCTTAAGAGCTTTAAACGGTTCAA 981
Db |||
QY 181 TyrSerAsnLeuValSerGlyAspGlyGlyThrIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTGCTTTCCAAAGAAATATCTGCTCAAGCTGATGGGAGCTTGT 1041
Db |||
QY 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTACCAAGTTTCTGCTGCTATGGCTTAACGAGGCTCTTATGCTTTGTAGCGCAAT 1101
Db |||
QY 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
QY 1102 GTTTCGAGGAGTAAGAGGGGAGGATTTGCTGCTGTTTCAGGATGGGACGAGGAGTCTCA 1161
Db |||
QY 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTTCCAGAAATATCTCGGTAGAGTTT 1221
Db |||
QY 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
```

QY 1222 GATGGAGAGTAGCCCGAGTAGGAGGAGATTATCTCCTACGGGAACGTTGCTTTCCTG 1281  
Db 281 AspGlyAenValAlaAargValGlyGlyGlyIlefySerfyrGlyAenValAlaPheLeu 300  
QY 1282 AATAATGAAACCTTTCTCAACAATGTTGCTTCTCCTGTTTACATTGCTGCTAAG 1341  
Db 301 AenAenGlyLyseThrLeuPheLeuAenAenValAlaSerProValfyrIleAlaAlaGlu 320  
QY 1342 CAACCAACAAAGTGGACAGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401  
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAenAenThrGlyAenGlyAlaIle 340  
QY 1402 TTCTGTAAAGATGGTGGCAA---GCAGATCCAATAAATCTCGGATCAGTTTCTTTGAT 1458  
Db 341 PheCylsAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAen 360  
QY 1459 CGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGTGTCTGGGAAGGGGAGCTATTAT 1518  
Db 361 GlyGluGlyValAlaPhePheSerSerAenValAlaAlaGlyLysGlyAlaIlefyr 380  
QY 1519 GCCAAAAAGCTCGGTTGCTAACTGTGGCCCTGTACAATTTTAAAGAAATATCGCTAAT 1578  
Db 381 AlalysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400  
QY 1579 GATGTGGAGCGAATTTATTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638  
Db 401 AspGlyGlyAlaIlefyLeuGlyLeuGlySerGlyGluLeuSerLeuSerAlaAenPyrGly 420  
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAACTGCCGATCTTAAT 1698  
Db 421 AspIleIlePheAenGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAenValAen 440  
QY 1699 GCGGTAACTGTCTCAACGCCATTTGATGGGATCGGAGGAGAAATAACGACATTA 1758  
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrLeu 460  
QY 1759 AGAGCTAAAGCAGGCGATCAGATTCTTTTAATGATCCCATCGAGATGCCAAACGGAAT 1818  
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAenProIleGluMetAlaAenGlyAen 480  
QY 1819 AACGAGCCGCGAGCTCTCCAACTTCTAAAAATTAACGATGGTGAAGGATACACAGG 1878  
Db 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAenGlyGluGlyfyrThrGly 500  
QY 1879 GATATTGTTTGTCTAAAGCAGTAGTCTTTGTACCAAAATGTTACGATAGCAAGGA 1938  
Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyfyrGlnAenValThrIleGluGlnGly 520  
QY 1939 AGGATTGTTCTTCGTGAAAGCAAAATPATCAGTGAATTCCTAAGTCAGACAGGTGG 1998  
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540  
QY 1999 AGTCTGTATATGAAGCTGGAGTACATGGGATTTGTATCTCCCAACCAACCAACAG 2058  
Db 541 SerLeuTyfyrMetGluAlaGlySerThrLeuAenPheValThrProGlnProGlnGln 560  
QY 2059 CCTCTCGCGCTAACTCAGTTGATCAGCTTTCATCTGCATTTGCTCTTCTCTTCTTG 2118  
Db 561 PropProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerLeu 580  
QY 2119 TTAGCAAAACAATGCAATTCAGATTCCTCTACCAATCTCCAGCGCAAGATTCATCCT 2178  
Db 581 LeuAlaAenAenAlaValThrAenProProThrAenProAlaGlnAenSerHisPro 600  
QY 2179 CGAGTCATTGGTAGCAACTGCTGCTTCTGTAGCAATTAGTGGGCTATCTTTTGGAG 2238  
Db 601 AlaValIleGlySerThrAlaGlyProValThrIleSerGlyProPhePheGlu 620  
QY 2239 GATTGTGATGATACGCTTATGATAGGTATGATGGCTAGGTTCTAATCAAAAAATCAAT 2298  
Db 621 AspLeuAenAenPheThrAlaTyfyrAenPyrAenPyrLeuGlySerAenGlnLysIleAen 640

QY 2299 GTCTCGAAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCATCAGATTTGACTCTA 2358  
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAenPheLeu 660  
QY 2359 GGGAAATGAGATGCTTAAGTATGGCTATCAAGAAAGCTGGAAGCTTGCCTGGATCCTAAT 2418  
Db 661 GlyAenGluMetProLysfyrGlyfyrGlnGlySerfyrPheLysLeuAlaIlePheAenPro 680  
QY 2419 ACAGCAATAATGTCCTTATCTCTGAAGACTACATGACTAAACCTGGGTATATCTCT 2478  
Db 681 ThrAlaAenAenGlyProTyfyrThrLeuLysAlaThrTrpThrLysThrGlyfyrAenPro 700  
QY 2479 GGGCTGAGCCAGTAGCTTCTTGGTCCAAATAGTTTATGGGATCCATTTAGATATA 2538  
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAenPhe 720  
QY 2539 CGATCTGGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATGTCGAGGATTA 2598  
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAenPheGlyAenGlyfyrCysArgGlyLeu 740  
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTTCTATCATGACCCGATGCTTTTAGTCAGGATAT 2658  
Db 741 TrpValSerGlyValSerAenPheSerfyrHisAenArgAenAlaLeuGlyGlnGlyfyr 760  
QY 2659 CGGTATATTAGTGGGTTTATCTCTTAGGACAACTCTACTTTGGATCATCGATGTTT 2718  
Db 761 ArgfyrIleSerGlyGlyfyrSerLeuGlyAlaAenSerfyrPheGlySerMetPhe 780  
QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTTAAAGATATATGATGTCGTTCCCAAT 2778  
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAenPyrValValCysArgSerAen 800  
QY 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCCCAACGCTTTATGATGATCCTAT 2838  
Db 801 HisHisAlaCysIleGlySerValfyrLeuSerThrLysGlnAlaLeuCysGlySerfyr 820  
QY 2839 TTGTTGGAGATCGTTTATCCGTCGTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898  
Db 821 LeuPheGlyAenAlaPheIleArgAlaSerfyrGlyPheGlyAenGlnHisMetLysThr 840  
QY 2899 TCATATACATTTGCAGAGGAGCGATGTTCTGTTGGGATAATACTGTCTGCTCGAGAG 2958  
Db 841 SerfyrThrPheAlaGluGluSerAenPheValArgTrpAenAenCysLeuValGlyGlu 860  
QY 2959 ATTGAGCGGATTAACGATTTGATTAATCCATCTAAAGCTCTATTGAATGATGCTGT 3018  
Db 861 IleGlyValGlyLeuProIleValThrThrProSerLysLeuTyfyrLeuAenGluLeuArg 880  
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGCGAT 3078  
Db 881 ProPheValGlnAlaGluPheSerfyrAlaAenPheHisGluSerPheThrGluGluGlyAen 900  
QY 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138  
Db 901 GlnAlaAenAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920  
QY 3139 TTTGATCAGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGCGCTTATCTGT 3198  
Db 921 PheAenArgCysSerSerThrHisProAenLysfyrSerPheMetGlyAlaTyfyrIleCys 940  
QY 3199 GATGTTATCGCACCATCTCTGGTACTGAGACAAACGCTCTTATCCATCAAGAGCATGG 3258  
Db 941 AspAlaTyfyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960  
QY 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGAGTGTGTTGTTAGAGGATCTATGATGCT 3318  
Db 961 ThrThrAenAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetfyrAla 980  
QY 3319 TCTTAAACAAGTAATATAGAAGTATATGGCCATGGAAGATATGAGTATCAGATGCTTCT 3378  
Db 981 SerLeuThrSerAenIleGluValfyrGlyHisGlyAenGlyfyrGluTyfyrArgAenThrSer 1000  
QY 3379 CGAGCTATGGTTTGTAGTGCAGGAAGTAGAGTCCGGTTC 3417



```
||||| 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
Db
RESULT 8
US-09-556-877-190
; Sequence 190, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probat, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yashir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469CS
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-190

Alignment Scores:
Pred. No.: 0 Length: 1006
Score: 5090.00 Matches: 979
Percent Similarity: 99.9% Conservative: 2
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 64.6% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-556-877-190 (1-1006)
QY 472 ATGGTTCCTCAAGGAATTTACGATGGGAGAGCTTAACGTGTATCATTTCCCTATACGTGT 531
Db
25 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 44
QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTTCTCGAGGAGAGATTAAACATTAAAAATCTT 591
Db
45 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64
QY 592 GACAAATCTATTGCGAGCTTGGCTTTAGTTGTTTGGGAACCTATTAGGGAGTTTACT 651
Db
65 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuGlySerPheThr 84
QY 652 GTTTTAGGAGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACAAATGGGCA 711
Db
85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104
QY 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGGGTTTTAAAGAAATATCC 771
Db
105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124
QY 772 TTTTCCAATGCAATTCATTACTTGGCGTACTGCTGCTGCGAACGACTAATAAGGGTAC 831
Db
125 PheSerAsnCyAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144
QY 832 CAGACTCCGACGACACATCTACCCGCTCTAATGGTACTATTATTCTAAACAGACTCTT 891
Db
145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 164
QY 892 TTGTACTCAATAGAGAGTTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGGA 951
Db
165 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 184
QY 952 GCTATAGAGCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGTTCTCCAGAA 1011
Db
185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCyValPheGlnGlu 204
QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCT 1071
||||| 205 AsnThrAlaGlnAlaAspGlyGlyAlaCyGlnValValThrSerPheSerAlaMetAla 224
1072 AACGAGGCTCCTATTTCCTTTGTAGCGAATCTTGAGAGTAAGAGGGGGAGGATGCT 1131
225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 244
1132 GCTGTTTCAGGATGGGACGAGGAGTGTTCATCTACTTCAACAGAAGATCCAGTAGTA 1191
245 AlaValGlnAspGlyGlnGlnGlyValSerSerThrThrGluAspProValVal 264
1192 AGTTTTTCCAGAAATATCTGCGGTAGATTTCATGGGAACCTAGCCGAGTAGGAGAGG 1251
265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 284
1252 ATTTTACTCTACGGGAGCTTCTCTGTAATATGAAAAACCTTGTTCTCAACCAAT 1311
285 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsn 304
1312 GTTGCTTCTCTGTTTACATTTGCTAAGCAACCAACAGTGGACAGCTTCTTAATAG 1371
305 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324
1372 AGTAATAATTACGGAGATGGAGGACTATCTTCTTAAGATGCTGCGCAACGAGATCC 1431
325 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344
1432 AATAACTCTGATCAGTTTCTTTCATGAGAGGAGTAGTTTCTTTAGTACCAATGTA 1491
345 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 364
1492 GCTGCTGGGAAAGGGGAGCTATTATTATGCCAAAAAGCTCTCGTTGCTAACTGTGCCCT 1551
365 AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCyGlyPro 384
1552 GTACAATTTTAAAGGAATATCGCTAATGATGTGAGCGATTTATTATTAGGAGAACTGGA 1611
385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlySerGly 404
1612 GAGCTCAGTTTCTGCTGATTTCGAGATATTATTTCGATGGGAATCTTAAAAAGAAC 1671
405 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424
1672 GCCAAGAGAAATCGCCGATGTTAATGCGTAACTGTGCTCTCAACGCCATTTCCGATG 1731
425 AlaLysGluAsnAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444
1732 GGATCGGAGGAGAAATAACGACATTAAAGAGCTAAAGCAGGCGCATCAGATTCTTTAAT 1791
445 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464
1792 GATCCATCGAGATCGCAACCGAAATAACCGCCAGCGAGCTTCCAAACTTCTAAAA 1851
465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484
1852 ATTAACGATGTGAAGGATACACAGGGATATTGTTTTGCTTAATGGAAGCAGTACTTTG 1911
485 IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504
1912 TACCAAAATGTTACGATAGAGCAAGAGGATTGTTCTTCGTGAAAGCGCAAAATTTATCA 1971
505 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 524
1972 GTGAATTTCTTAAGTCAGACAGTGGAGTCTGTATATCGAGCTTCCAACTTCTGGAT 2031
525 ValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 544
2032 TTTTGAATCTCCCAACCAACCAACAGCTCTCGCGCTTAATCAGTGTGATCAGCTTTCC 2091
545 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 564
2092 AATCTGCAATTTGCTCTTTCTTTGTTAGCAAAACAATGCAGTTACGAATCTCTTACC 2151
```



Db 565 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 584  
Qy 2152 AATCCTCCAGCGCAAGATTCTCTCTCGAGTCAATGGTAGCACAACTGCTGGTTCGTT 2211  
Db 585 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrAlaGlySerVal 604  
Qy 2212 ACAATTAGTGGGCTATCTTTTGGAGATTGGATGATGATACAGCTTATGATAGTAGAT 2271  
Db 605 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 624  
Qy 2272 TGGCTAGGTTCTAATCAAAAATCAATGCTGGAATTCAGATTAGGAGCTTAGGCCCCCA 2331  
Db 625 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 644  
Qy 2332 GCTAATCCCCCATCAGATTGTGACTTAGGGAATCAGATGCTAAGTAGTGGCTATCAAGGA 2391  
Db 645 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyThrGlnGly 664  
Qy 2392 AGCTGGAAGCTGCGTGGGATCCTTAATACAGCAAAATATGGTCTCTTATCTCTGAAAGCT 2451  
Db 665 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 684  
Qy 2452 ACATGGACTAAACTGGGTATATCCTGGGCTCAGCGAGTAGCTCTTTGGTTCCAAAT 2511  
Db 685 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 704  
Qy 2512 AGTTTATGGGATCCATTTAGATATACATCTCGGCATTCAGCAATTCAGCAAGTGTG 2571  
Db 705 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724  
Qy 2572 GATGGGCGCTTATATGTCGAGGATTATGGGTTCTGGAGTTTCGAATTTCTTATCAT 2631  
Db 725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPheTyrHis 744  
Qy 2632 GACCGGATGTTTAGTCAGGATATCGGTATATTAGTGGGGTTATTCCTTAGGACA 2691  
Db 745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764  
Qy 2692 AACTCCTACTTTGGATCATCCATGTTGGTCTAGCATTTACGAAGTATTTGGTAGATCT 2751  
Db 765 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784  
Qy 2752 AAAGATTATGATGTGTGTTCCAAATCATCATGCTTCATAGGATCCGTTTATCTATCT 2811  
Db 785 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 804  
Qy 2812 ACCCAACAGCTTTATGATGATCCTATTGTTGCGAGATGCTTTATCCGTGCTAGCTAC 2871  
Db 805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824  
Qy 2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCAGAGAGAGGATGTTGCT 2931  
Db 825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 844  
Qy 2932 TGGGATAATAACTGTCTGGCTGGAGAGATTGGAGCGGATTAACCAATTGTGATTACTCA 2991  
Db 845 TrpAspAsnAsnCysLeuAlaGlyIleGlyAlaGlyLeuProIleValIleThrPro 864  
Qy 2992 TCTAAGCTCTATTGAAAGATTGGTCTTCTTCGTCGAGCTGAGCTTCTTATCCGAT 3051  
Db 865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884  
Qy 3052 CATGAATCTTTTACAGAGAGGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCCTTA 3111  
Db 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904  
Qy 3112 AATCTATCAGTTCTCTGTTGGAGTGAAGTTGATCGATGCTTCTAGTACATCCTAATAA 3171  
Db 905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924  
Qy 3172 TATAGCTTTATCGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231  
Db 925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944

Qy 3232 ACCTCTCTATCCCATCAAGAGACATCGACCAACAGATGCTTTTCATTTAGCAAGACATGGA 3291  
Db 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964  
Qy 3292 GTTGTGTTAGAGATCTATGTATGTTCTCTCAACAAGTAATATAGAAGTATATGGCCAT 3351  
Db 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984  
Qy 3352 GGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTGTGATGCGAAGTAGAGTC 3411  
Db 985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004  
Qy 3412 CGGTTT 3417  
Db 1005 ArgPhe 1006  
RESULT 9  
US-09-620-412C-190  
; Sequence 190, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620.412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 190  
; LENGTH: 1006  
; TYPE: PR1  
; ORGANISM: Chlamydia  
US-09-620-412C-190  
Alignment Scores:  
Pred. No.: 0 Length: 1006  
Score: 5090.00 Matches: 979  
Percent Similarity: 99.9% Conservative: 2  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 64.6% Indels: 0  
Gaps: 0  
DB: 2  
US-10-701-844-1 (1-4435) x US-09-620-412C-190 (1-1006)  
Qy 472 ATGTTCTCTCAAGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATCTGTT 531  
Db 25 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 44  
Qy 532 ATAGAGATCCGAGTGGGACTACTGTTTCTGAGAGAGATTACATTTAAAAATCTT 591  
Db 45 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64  
Qy 592 GACAAATCTATTCAGCTTTCCTTAAAGTTGTTTGGAACTTATTAGGAGTTTACT 651  
Db 65 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84  
Qy 652 GTTTTAGGAGAGGACACTCGTTGACTTTCAGAAACATACCGACTTCTACAAATGGGCA 711  
Db 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104  
Qy 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGGGTTTAAAGAATTATCC 771  
Db 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124  
Qy 772 TTTTCCAAATTCGAATTCATTACTTCCCTGCTGCTGCAACGACTAATAAGGATGAC 831  
Db 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144  
Qy 832 CAGACTCCGACGACCAACATCTACACCGTCTAATGCTACTATTATTCTAAAAACAGATCTT 891  
Db 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 164

QY	892	TTGTTACTCAATAATGAGAAGTTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGGA	951	QY	1972	GTGAATTTCTTAAGTTCAGACAGTGGGAGTCTGTATATAGGAAGCTGGGAGTACATGGGAT	2031
DB	165	LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly	184	DB	525	ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp	544
QY	952	GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGTTCTCTCCAGAA	1011	QY	2032	TTTGTAATCTCCAAACCCACCAACAGCTCTCTCGGCTTAATCAGTGTATCAGCTTCC	2091
DB	185	AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu	204	DB	545	PheValThrProGlnProGlnProAlaAlaAsnGlnLeuIleThrLeuSer	564
QY	1012	AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCT	1071	QY	2092	AATCTGCATTTCTCTTCTTCTTTGTAGCAAAACAATGCAGTTTACGAATCCTCTTACC	2151
DB	205	AsnThrAlaGlnAlaAspGlyAlaCysGlnValValThrSerPheSerAlaMetAla	224	DB	565	AsnLeuHleLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr	584
QY	1072	AACGAGCTCTTATTCCTTCTGAGCAATGTTGCAGGAGTAGAGGGGAGGAGTTGCT	1131	QY	2152	AATCTCCAGCCCAAGATTTCTCATCTGCAGTCATTTGGTAGCACAACTCTGCTGTTCTGT	2211
DB	225	AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla	244	DB	585	AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal	604
QY	1132	GCTGTTTCAGATGGGAGGAGGTGTCATCATCTTCAACAGAAGATCCAGTAGTA	1191	QY	2212	ACAATTAGTGGGCTTATCTTTTGTAGGATTTGGATGTATACAGCTTATCATATAGTATGAT	2271
DB	245	AlaValGlnAspGlyGlnGlnValSerSerSerThrThrGlnAspProValVal	264	DB	605	ThrIleSerGlyProIlePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp	624
QY	1192	AGTTTTTCCAGAAATCTCGGTAGATGTTGTAGGAAAGTACCCGAGTAGGAGGG	1251	QY	2272	TGCTAGGTTCTTAATCAAAAAATCAATGTCTCTGAAATTAACAGTTAGGGACTAAGCCCCCA	2331
DB	265	SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly	284	DB	625	TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro	644
QY	1252	ATTACTCTTACGGGAACGTTGCTTCTCCTGAATAATGGAAAACTTTGTTCTCAACAT	1311	QY	2332	GCTAATGCCCCATCAGATTTGACTCTAGGAAATGAGATGCCCTAAGTATGCTATCAAGGA	2391
DB	285	IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn	304	DB	645	AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyThrGlnGly	664
QY	1312	GTTGCTTCTCCTGTTTACATGCTGTAGCAACCAACAGTGCACAGGCTTCTAATACG	1371	QY	2392	AGCTGAAAGCTTGGTGGGATCCTTAATACAGCAAAATAATGGTCTTATCTCTGAAAGCT	2451
DB	305	ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr	324	DB	665	SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla	684
QY	1372	AGTAATAATACCGAGATGAGGAGCTATCTTCTGTAAGAATGTCGCGAAGCAGATCC	1431	QY	2452	ACATGCACTAAACCTGGGTATATCTCGGCGCTGACGAGTAGCTTCTTCTTGGTTCCAAAT	2511
DB	325	SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer	344	DB	685	ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn	704
QY	1432	AATACTCTGGATCAGTTCTTGTAGGAGAGGAGTAGTTTCTTTAGTACGAATGTA	1491	QY	2512	AGTTTATGGGATCCATTTTAGATATACGATCTCGCATTTACAGCAATTCAGCAAGGTG	2571
DB	345	AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal	364	DB	705	SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal	724
QY	1492	GCTGCTGGGAAAGGGAGCTATTATGCCCCAAAGCTCTCGTTGCTAACTGTGGCCCT	1551	QY	2572	GATGGGCTCTTATTGTGAGGATTTAGGTTTCTGGAGTTTCGAATTTCTCTCATCAT	2631
DB	365	AlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro	384	DB	725	AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis	744
QY	1552	GTACAAATTTTAAGGAATATCGCTAATGATGGTGGACGATTTATTATGAGGAATCTGGA	1611	QY	2632	GACCGGATCTTTAGTCTAGGATATCGGTATATTAGTGGGGGTATTCTCTTAGGAGCA	2691
DB	385	ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGly	404	DB	745	AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla	764
QY	1612	GAGCTCAGTTTATCTGCTGATTATGAGATATTATTTTCGATGGGATCTTAAAGAAC	1671	QY	2692	AACTCTACTTTTGGATCATCGATGTTTGGTCTTAGCATTTTACCGAAGTATTGTTAGTATCT	2751
DB	405	GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr	424	DB	765	AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer	784
QY	1672	GCCAAAGAGATCTCGCGATGTTAATGGCGTAACTGTGCTCACAAGCCATTTCGATG	1731	QY	2752	AAAGATTATGATGTGTGCTTCCAAATCATCATCTGCTTAGCATAGATCCGTTTATCTATCT	2811
DB	425	AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet	444	DB	785	LysAspTyrValValCysArgSerAsnHisHleAlaCysIleGlySerValTyrLeuSer	804
QY	1732	GGATCGGAGGGAATAACACATTAAGAGCTAAAGCAGGCGCATCAGATTCTCTTTAAT	1791	QY	2812	ACCCAACAGCTTTATGTGATCTTATTTCTCGGAGATCGGTTTATCCGTGCTAGCTAC	2871
DB	445	GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn	464	DB	805	ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr	824
QY	1792	GATCCCCATCGAGATGGCAAAACGGAATAACACAGCCAGCGCATCTTCCAAACTCTTAAA	1851	QY	2872	GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTTCAGAGGAGAGCGATGTTCTGT	2931
DB	465	AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys	484	DB	825	GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg	844
QY	1852	ATTAACCATGCTGAGGATACACAGGAGTATCTTTTGTCTAATGGAGCAGTACTTTG	1911	QY	2932	TGGGATAATAACTGCTCTGCTGGAGAGATTGGAGCGGATTTACCGATTGTGATTACTCCA	2991
DB	485	IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerThrLeu	504	DB	845	TipAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro	864
QY	1912	TACCAAAATGTTACGATAGACAGGAGGATGTTCTTCTCGTGAAGGCGCAAAATTTATCA	1971	QY	2992	TCTAAGCTCTATTGATGATGTCGCTCTTCTCGTCAAGCTGAGTTTCTTATGCCGAT	3051
DB	505	TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer	524	DB	865	SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp	884
				QY	3052	CATGAATCTTTTACAGAGGAAGGCGATCAAGCTCGGGCATTTCAAGAGGAGCATCTCTCTA	3111

Db 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904  
Qy 3112 AATCTATCAGTTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCCTAATAAA 3171  
Db 905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924  
Qy 3172 TATAGCTTTATGGCGCTTATATCTGATCTGATGCTTATCGCACATCTCTGGTACTGAGACA 3231  
Db 925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944  
Qy 3232 ACGTCTCTATCCCATCAAGACATGGACACAGATGCTTTCATTATAGCAAGACATGGA 3291  
Db 945 ThrLeuLeuSerHisGlnGluThrTrpThrAspAlaPheHisLeuAlaArgHisGly 964  
Qy 3292 GTTGTGGTTAGAGATCTATGATGCTTCTCTAAACAAGTAATATATAGATATATGGCCAT 3351  
Db 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984  
Qy 3352 GGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTGAGTGCAGGAGTAGATC 3411  
Db 985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004  
Qy 3412 CGGTTC 3417  
Db 1005 ArgPhe 1006

## RESULT 10

US-09-598-419-190

Sequence 190: Application US/09598419

Patent No. 6565856

GENERAL INFORMATION:

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Scholler, John

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND

TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.469C6

CURRENT APPLICATION NUMBER: US/09/598,419

CURRENT FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 357

SOFTWARE: FastSeq for Windows Version 3.0/4.0

SEQ ID NO 190

LENGTH: 1006

TYPE: PRT

ORGANISM: Chlamydia

US-09-598-419-190

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	5090.00	1006	979
Percent Similarity:	99.9%	Conservative:	2
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	64.6%	Indels:	0
DB:	2	Gaps:	0

US-10-701-844-1 (1-4435) x US-09-598-419-190 (1-1006)

Qy 472 ATGTTCTCTCAAGGAATTTACGATGGGAGAGCTTAACCTGTATCATTTCCCTATATCTGTT 531  
Db 25 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 44  
Qy 532 ATAGAGATCCGAGTGGGACTACTGTTTTTCTCGAGGAGAGTTTAACTATAAAATCTT 591  
Db 45 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64  
Qy 592 GCAATCTTATTCGAGCTTGGCTTAACTTGTGGAACCTTATAGGGAGTTTACT 651  
Db 65 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84  
Qy 652 GTTTTAGGAGAGGACACTCGTTGACTTTCGAGAACATACGCACTTCTACAAATGGGCA 711  
Db 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104

Qy 712 GCTCTAAGTAATAGCGCTGCTGATGAGTCTTTACTATTGAGGCTTTTAAAGATTATCC 771  
Db 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124  
Qy 772 TTTTCCAAATTCGAATTCATTACTTTCGCGTCTGCTGCTGCAACGACATAATAAGGGTAGC 831  
Db 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144  
Qy 832 CAGACTCCGACGACAAACATCTACACCGTCTAATCGTACTATTATTCTTAAACAGATCTT 891  
Db 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 164  
Qy 892 TTGTACTCATATAATAGAGTCTCTCATCTCTATAGTAATTTAGTCTCTGGAGATGGGGA 951  
Db 165 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 184  
Qy 952 GCTATAGATGCTAAGAGCTTAACCGTTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAAGAA 1011  
Db 185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 204  
Qy 1012 AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCACAGTCTTCTCTGCTATGGCT 1071  
Db 205 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224  
Qy 1072 AACGAGGCTCTATTTGCTTGTAGCGAATCTTCAGAGTAAAGAGGGGGAGGATTCCT 1131  
Db 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 244  
Qy 1132 GCTGTTTCAGGATGGGACGAGGAGTGTCACTCATCTCTCAACAGAAGATCCAGTAGTA 1191  
Db 245 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal 264  
Qy 1192 AGTTTTTCCAGAAATACCTCGGTAGAGTTTGATGGGAACGTAGCCCGAGTAGGAGAGGG 1251  
Db 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly 284  
Qy 1252 ATTTACTCTACGGGAGGAGTTCCTTCTGTAATATGAAAAACCTTGTCTCAACAAAT 1311  
Db 285 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 304  
Qy 1312 GTTGCTTCTCTGTTTACATTCGTAAAGCAACCAACAGTGGACAGCGCTTCTTAATACG 1371  
Db 305 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324  
Qy 1372 AGTAATAATACGAGATGGAGGAGCTATCTCTGTAAGAATGTCGCAAGCAGGATCC 1431  
Db 325 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344  
Qy 1432 AATAACTCTGGATCAGTTTCTCTTGTATCGAGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
Db 345 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 364  
Qy 1492 GCTGCTGGAAAGGGGAGCTATTATTCGCAAAAAGCTCTCGGTTGTCTAACTGTGCGCCT 1551  
Db 365 AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 384  
Qy 1552 GTACAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGATTATTTTAGGAGAATCTGA 1611  
Db 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGly 404  
Qy 1612 GAGCTCAGTTTATCTGCTGATATCGAGATATTATTTTCGATGGGAATCTTAAAGACA 1671  
Db 405 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424  
Qy 1672 GCCAAGAGAAATGCTGCGCATGTTAATCGCGTAACGTGTGTCTCAAGGCCATTTCCGATG 1731  
Db 425 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444  
Qy 1732 GGATCGGAGGGAAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATCTCTTTAAT 1791  
Db 445 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464  
Qy 1792 GATCCCATCGAGATGGCAACGGAATAACCGCCAGCGCAGCTCTTCCAAACTTCTTAAAA 1851

465	AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerIysLeuLeuIys	484
1852	ATTAAACGATGGTGAAGGATACACAGGGGATATTGTTTTGCTTAATGAAGACGACTGTTG	1911
485	IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu	504
1912	TACCAAAATGTTACGATAGACAGCAAGGATGTTCTTCGTCAAAAGGCAAAATATATCA	1971
505	TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluIysAlaIysLeuSer	524
1972	GTGAATTCTTAAGTCACAGACGATGGGAGTCTGTATATGAAGCTGGGATCATCGGAT	2031
525	ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp	544
2032	TTTTGTAACTCCAAACCACACACACGCTCCCTCCGCTTAATCAGTTGATCAGCTTTCC	2091
545	PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer	564
2092	AATCTGATTTGTCCTCTTTCTCTTTGTTAGCAAAACAATCAGTTAGCAATCTCTACCC	2151
565	AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr	584
2152	AATCCTCCAGCGCAAGATTCTATCTCTGTCAGTCATTGTTGATAGCAACACTGCTGTTCTGTT	2211
585	AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal	604
2212	ACAATTAGTGGGCTATCTTTTTTGAGATTTCGATGATACAGCTTATGATAGTATGAT	2271
605	ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp	624
2272	TGGCTAGGTTCTAATCAAAAAATCAATGTCCTCGAAATTACAGTTAGGACCTAAGCCCCCA	2331
625	TrpLeuGlySerAsnGlnIysIleAsnValLeuIysLeuGlnLeuGlyThrIysProPro	644
2332	GCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCTTAAGTATGCTATCAAGGA	2391
645	AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProIysTyrGlyTyrGlnGly	664
2392	AGCTGGAAGCTTGGCTGGGATCTTAATACAGCAAAATATGGCTCTTATCTCTGGAAGCT	2451
665	SerTrpIysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuIysAla	684
2452	ACATGGACTAAACCTGGGTATACTCTGGGCTCAGCGCAGTAGCTTCTTTGGTTCCAAAT	2511
685	ThrTrpThrIysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn	704
2512	AGTTTATGGGGATCCATTTTATAGATATACGATCTCGGCATTACGCAATTCAGCAAGTGTG	2571
705	SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal	724
2572	GATGGCGCTCTTATTTCGAGGATATATGGGTTTCTGGAGTTTCGAAATTTCTTCTATCAT	2631
725	AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis	744
2632	GACCGCATGCTTTTAGGTCAGGATATCGGTATATTAGTGGGGTATTCTCTAGGACCA	2691
745	AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla	764
2692	AACTCCTATTGGGATCATCGATTTGGTCTAGCATTTACGAAAGTATTTGGTAGATCT	2751
765	AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer	784
2752	AAAGATTATGATGTGTGTCGTTCCGATCATCATCTGTCATAGGATCCGTTTATCTATCT	2811
785	IysAspTyrValValCysArgSerAsnHisIleAlaCysIleGlySerValTyrLeuSer	804
2812	ACCCAAACAGCTTTATGTGGATCCCTATTGTTTCGGAGATCGGTTTATCCGTGCTAGCTAC	2871
805	ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr	824
2872	GGGTTTGGGAATCAGCATATCAAAACCTCATATACATTTGCAGAGGACGATGTTCTGT	2931

825	GlyPheGlyAsnGlnHisMetIysThrSerTyrThrPheAlaGluSerAspValArg	844
2932	TGGGATAATAAATCTGCTGGCTGGAGAGATGGAGCGGGATTACCGATTGTGATTACTCCA	2991
845	TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro	864
2992	TCATAAGCTCTATTGTAATGAGTTGCCTCTTTTCGTCGAAGCTGAGTTTCTTATGCCGAT	3051
865	SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp	884
3052	CATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGGCATTCGAAGCGGACATCTCCCTA	3111
885	HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheIysSerGlyHisLeuLeu	904
3112	AATCTATCAGTTCCTCTGTGGAGTGAAGTTTGATTCGATGTTTCTAGTACACATCCTAATAA	3171
905	AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys	924
3172	TATAGCTTTATGGCGGCTTATATCTGTGATGCTTTATCGACCAATCTCTGGTACTGAGACA	3231
925	TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr	944
3232	ACGCTCCTATCCCATCAAGAGACATCGACAACAGATGCCCTTCATTTAGCAAGACATGGA	3291
945	ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly	964
3292	GTTCGTGTTAGAGGATCTATGATGCTTCTCTAACAGTAAATATAGAAGTATATGGCCAT	3351
965	ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis	984
3352	GGAAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGGTGCAGGAAGTAGAGTC	3411
985	GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal	1004
3412	CGGTTC	3417
1005	ArgPhe	1006

```

RESULT 11
US-09-556-877-176
; Sequence 176, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-556-877-176

```

Alignment Scores:		
Pred. No.:	0	982
Score:	504.00	978
Percent Similarity:	99.8%	
Best Local Similarity:	99.6%	Conservative: 2
Query Match:	64.5%	Mismatches: 2
DB:	2	IndeIs: 0
		Gaps: 0

US-10-701-844-1 (1-4435) x US-09-556-877-176 (1-982)

QY 472 ATGGTTCTCAAGGAATTACATGGGAGACGTTAACTGTATATCAATTTCCCTATCTGTT 531  
Db : : : : :  
QY 1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20  
QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTCGCGAGGAGGTTAACTAAATAATCTT 591  
Db : : : : :  
QY 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40  
QY 592 GACAAATCTATTGACGCTTGCCTTAAAGTTGTTTGGAACTATTAGGAGGTTTACT 651  
Db : : : : :  
QY 41 AsnAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60  
QY 652 GTTTTAGGAGAGGACACTCGTTGACTTTCGAGAAACATACGACACTTCTACAAATGGGCA 711  
Db : : : : :  
QY 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80  
QY 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGGTTTAAAGAAATATCC 771  
Db : : : : :  
QY 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 100  
QY 772 TTTTCCAAATGCAATTCATTACTTCCGCTACTGCTGCTGCAACGACTAAATAAGGTAGC 831  
Db : : : : :  
QY 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 120  
QY 832 CAGACTCCGACGACAACTACTACCCGCTCTAATGGTACTATTATTCTAAACACAGATCTT 891  
Db : : : : :  
QY 121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 140  
QY 892 TTGTTACTCAATAATAGAGAGTTCTCATCTATAGTAAATTAGTCTCTGGAGATGGGGA 951  
Db : : : : :  
QY 141 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160  
QY 952 GCTATAGATGCTAAGAGCTTCAAGGTTCAAGAAATTTAGCAAGCTTTGTCTTCCAAGAA 1011  
Db : : : : :  
QY 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180  
QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACAGATTTCTCTGCTATGGCT 1071  
Db : : : : :  
QY 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200  
QY 1072 AACGAGGCTCTATTGCTTGTAGCGAATGTTGACGAGTAAAGAGGGAGGATGCT 1131  
Db : : : : :  
QY 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 220  
QY 1132 GCTGTTCCAGATGGGCGAGGAGTGCATCATCTTCAACAGAGATCCAGTAGTA 1191  
Db : : : : :  
QY 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrSerThrGluAspProValVal 240  
QY 1192 AGTTTTTCCAGAAATACTCGGCTAGAGTTTGTAGGAACTAGCCCGAGTAGGAGAGG 1251  
Db : : : : :  
QY 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 260  
QY 1252 ATTTACTCTACGGAACTGCTTTCCTGGAATAATATGGAAAAACCTTGTTCTCAACAAT 1311  
Db : : : : :  
QY 261 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 280  
QY 1312 GTTGCTTCTCCTGTTTACATTCGCTAAGCAACCAAGTGGACAGGCTTCTAATACG 1371  
Db : : : : :  
QY 281 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300  
QY 1372 AGTAATAATTACGGAGATGGAGGACTATCTTCTGTAAGAATGGTCGCAAGCAGGATCC 1431  
Db : : : : :  
QY 301 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320  
QY 1432 AATAACTCTGGATCAGTTCTTGTGATGGAGGGAGTAGTTTCTTTAGTAGCAATGTA 1491  
Db : : : : :  
QY 321 AsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSerAsnVal 340  
QY 1492 GCTGCTGGGAAAGGGAGCTATTATGCAAAAAGCTCTCGTTGCTAACTGTGGCCCT 1551  
Db : : : : :  
QY 341 AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 360

QY 1552 GTACAATTTTAAAGGAATATCGCTAATGATGGTCGAGCGATTATTATTAGGAGAACTCGA 1611  
Db 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlySerGly 380  
QY 1612 GAGCTCAGTTTATCTGCTGATTATGGAGATATATTTTCGATGGAAATCTTAAAGAACAA 1671  
Db 381 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400  
QY 1672 GCCAAGAGAAATGCTGCCGATGTTAATGGCGTAATGTGTCTCTCAACGCCAATTTCCGATG 1731  
Db 401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420  
QY 1732 GGATCGGAGGAGAAAATAACGACATTAAGACTAAACGAGCGCATCAGATTCTCTTTAAT 1791  
Db 421 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440  
QY 1792 GATCCCATCGAGATGGCAACCGAAATAACAGCCAGCGCAGTCTTCCAAACTTCTAAAA 1851  
Db 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 460  
QY 1852 ATTAACGATGTTGAAGGATACACAGGGGATATTGTTTTCCTAATGGAAGCAGTACTTTG 1911  
Db 461 IleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480  
QY 1912 TACCAAAATGTCATAGACCAAGAGGATGTTGTTCTTCTGTAAGAACGCAAAATATCA 1971  
Db 481 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500  
QY 1972 GTGAATTCCTTAAGTCAGACAGGTGGGAGTCTGTATATATGGAAGCTGGCAGGTACATGGAT 2031  
Db 501 ValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 520  
QY 2032 TTTGTAACTCCCAACACCAACAGCCTCTCGCGCTAATCACTAGTTGATCAACGCTTTC 2091  
Db 521 PheValThrProGlnProGlnProAlaAlaAsnGlnLeuIleThrLeuSer 540  
QY 2092 AATCTGCAATTTGCTCTTCTTCTTCTTGTAGCAAAACAATGACAGTTACGAATCTCTTACC 2151  
Db 541 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560  
QY 2152 AATCCTCCAGCGCAAGATTCTCATCTCGAGTCATTGTTAGCACAACTGCTGGTTCTGTT 2211  
Db 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580  
QY 2212 ACAATTAGTGGCCCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGGTATGAT 2271  
Db 581 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 600  
QY 2272 TGGCTAGGTTCTAATCAAAAAATCAATGCTCTGAAATTTACAGTTAGGCACTAAGCCCCA 2331  
Db 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620  
QY 2332 GCTAATGCCCATCAGATTTGACTCTAGCGAATGAGATGCCCTAAGTATGGCTATCAAGA 2391  
Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyGlnGly 640  
QY 2392 AGCTGGAAGCTTTCGCTGGGATCCTAATACAGCAAAATAATGGTCTTATATCTCTGAAGCT 2451  
Db 641 SerTyrLysLeuAlaTyrAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660  
QY 2452 ACATGGACTAAACCTGGGTATATCTCGGCTGAGCGAGTAGTCTCTTGGTTCCAAAT 2511  
Db 661 ThrTyrThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680  
QY 2512 AGTTTATCGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATTCACAGAGTGG 2571  
Db 681 SerLeuTyrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700  
QY 2572 GATGGGCGCTCTATTGTGCGAGGATATTGGGTTCTCGGAGTTTCGAATTTCTTCTATCAT 2631  
Db 701 AspGlyArgSerTyrCysArgGlyLeuTyrValSerGlyValSerAsnPhePheTyrHis 720

2632	QY	GACCGCATCTTTAGTGTACAGGATATCGGTATATATAGTCGGGGTTATTCTCTTAGAGCA	2691
721	Db	AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla	740
2692	QY	AACTCCCTACTTTGGATCATCGATGTTTGGTCTAGCATTTAACCAAGTATTTGCTAGATCT	2751
741	Db	AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer	760
2752	QY	AAAGATATATGATGTGTCTGTTCCAATCATCATGCTTGTGCATAGATCCGTTTATCTATCT	2811
761	Db	LysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyrLeuSer	780
2812	QY	ACCCAAACAGCTTATCTGTGATCTCTATTTCTTCGGAGATCGTTTATCCGTGCTAGCTAC	2871
781	Db	ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr	800
2872	QY	GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTTCGAGAGGAGACCGATGTTCGT	2931
801	Db	GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg	820
2932	QY	TGGGATTAATACTGCTCTGGCTGGAGAGATTGGAGCGGGATTACCGATTGTGATTACTCCA	2991
821	Db	TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro	840
2992	QY	TCTAAGCTCTATTTGAATGAGTTCCGTCCTTTCGTGCAAGCTGAGTTTCTTATCCCGAT	3051
841	Db	SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp	860
3052	QY	CATGAATCTTTTACAGAGGAGGCGGATCAAGCTCGGCATTTCAAGACGGACATCTCCTA	3111
861	Db	HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu	880
3112	QY	AATCTATCAGTTCCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTAATAAA	3171
881	Db	AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys	900
3172	QY	TATAGCTTTATGGCGGCTTATATCTGTGATCTTATCGCACCATCTCTGGTACTGAGACA	3231
901	Db	TyrSerPheMetAlaIaIaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr	920
3232	QY	ACGCTCTCTATCCCATCAAGACACATGGACAAACAGATCGCTTTCATTTAGCAAGACATGGA	3291
921	Db	ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly	940
3292	QY	GTTCTGTTAGAGGATCTATGATGCTTCTCTCAACAGTATATAGAGTATATGGCCAT	3351
941	Db	ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis	960
3352	QY	GGAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTGAAGTCGAGGAGTAGATC	3411
961	Db	GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal	980
3412	QY	CGGTTTC	3417
981	Db	***Phe	982

**RESULT 12**

US-09-620-412C-176  
; Sequence 176, Application US/09620412C

Patent No. 6448234

**GENERAL INFORMATION:**

APPLICANT: Steven P. Fling

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

**; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION**

**FILE REFERENCE: 210121.469**

; CURRENT APPLICATION NUMBER: US/09/620,412C

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 363

```

; SOFTWARE: FastSEQ for Windows Version 3.0/4.0

```

; SEQ ID NO 176  
T TACCTG 092

```

; LENGTH: 982
TYPE: COM

```

TYPE: PKT  
OPERATING: CHLAMPIDIA

ORGANISM: Chlamydia

```

; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-412C-176

```

	Alignment Scores:		
Pred. No.:	0	Length:	982
Score:	5084.00	Matches:	978
Percent Similarity:	99.8%	Conservative:	2
Best Local Similarity:	99.6%	Mismatches:	2
Query Match:	64.5%	Indels:	0
DB:	2	Gaps:	0

U.S.-10-701-844-1 (1-4435) x U.S.-09-620-412C-176 (1-982)

QY	472	ATGTTCTCTCAAGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATACTGTT	531
DB	1	MetileproGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal	20
QY	532	ATAGGAGATCCGAGTGGGACTACTGTTTTTTCTGCAGGAGAGTTAAACAATAAAAAATCTT	591
DB	21	IleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu	40
QY	592	GACAAATTTCTATTGCAGCTTTGCCCTTTAAGTTGTTTTTGGGAACCTATTATTAGGAGTTTTACT	651
DB	41	AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr	60
QY	652	GTTTTAGGAGAGACACTCGTTGCATCTTCGAGAACATACGGACTTCTACAAATGGGSCA	711
DB	61	ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla	80
QY	712	GCTCTAAGTAAATACGGCTGCTGATGGACTGTTTACTATTATTCAGGGTTTTAAAGAATATCC	771
DB	81	AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer	100
QY	772	TTTTCCAAATTCGAATTCATTACTTTCGCGTACTGCTGCTGCAACGACTAATAAGGGTAGC	831
DB	101	PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaIaThrThrAsnLysGlySer	120
QY	832	CAGACTCCGACGACAAACATCTACACCGCTCAATGGTACTATTATTATCTTAAACACAGACTTT	891
DB	121	GlnThrProThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu	140
QY	892	TGTGTACTCAATAATGAGAAGTCTCATCTTATAGTAAATTTAGTCTCTGGAGATGGGGGA	951
DB	141	LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly	160
QY	952	GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTACAGACTTTGTGTCTTCCAAGAA	1013
DB	161	AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu	180
QY	1012	AATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCCACAGTTCTCTGCTATGGCT	1071
DB	181	AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla	200
QY	1072	AACGAGGCTCTATTGCTTGTAGCCGATCTTCGCGAGTAAGAGGGGGGAGGATTCCT	1131
DB	201	AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla	220
QY	1132	GCTGTTTCAGGATGGGCACGAGGAGTGTCATCATCTACTTCAACAGAAGATCCAGTAGTA	1191
DB	221	AlaValGlnAspGlyGlnGlnGlyValSerSerThrSerThrGluAspProValVal	240
QY	1192	AGTTTTTCCAGAAATATCGCGGTAGAGTTTGATGGGAACGTAGCCCGAGTAGGACGAGGG	1251
DB	241	SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly	260
QY	1252	ATTTTACTCTTCGCGGAACGTTGCTTTCTCTGAATAATGGAACAACTGTGTTCTCAACAAT	1311
DB	261	IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn	280
QY	1312	GTGTGCTTCTCTGTTTACATTGCTGCTAAGCAACCAACAGTGGGACAGGGTCTTAATAGC	1371

Db 281 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300  
QY 1372 AGTAATAATTACGGAGATGGAGGAGCTATCTTCTGAAGAATGGTGGCCAAAGCAGGATCC 1431  
Db 301 SerAsnAsnTyrGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320  
QY 1432 AATAACTCTGATCAGTTCCTTTGATCGAGAGGAGTAGTTCCTTTCTTAGTACCAATGTA 1491  
Db 321 AsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSerAsnVal 340  
QY 1492 GCTCTGGGAAGGGGAGCTATTATGCCAAAGAGCTCGGTGCTAACTGGCCCT 1551  
Db 341 AlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 360  
QY 1552 GTACAAATTTTAAAGAAATCGCTAATGATGTGGAGCGATTATTTAGGAGAACTCTGGA 1611  
Db 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlySerGly 380  
QY 1612 GAGCTCAGTTCCTGATATGAGAGATATATTTTCGATGGGAATCTTTAAAGAACA 1671  
Db 381 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400  
QY 1672 GCCAAGAGAGATGCTCCGATGTTAATCGCGTAAGTGTCTCTCAAGCCATTTCCGATG 1731  
Db 401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420  
QY 1732 GGATCGGAGGAGAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATCTCTTTAAT 1791  
Db 421 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440  
QY 1792 GATCCATCGAGATGGCAACGGAATAAACCCAGCCAGCGCAGTCTTCCAAACTCTTAAAA 1851  
Db 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 460  
QY 1852 ATTAACGATGGTGAAGGATACACAGCGATATGTTTTGCTTAATGGAGCGATCTTTG 1911  
Db 461 IleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480  
QY 1912 TACCAAAATGTTACCATAGACAGCAAGGATGTTCTTGTGAAAGGCAAAATATATCA 1971  
Db 481 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGlyLysAlaLysLeuSer 500  
QY 1972 GTGAATCTCTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGGAGTACATGGAT 2031  
Db 501 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 520  
QY 2032 TTTGTAACTCCACACCAACCAACAGCTCTCGCGCTAATCAGTTGATCAGCTTCC 2091  
Db 521 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 540  
QY 2092 AATCTGCATTTCTCTTCTTCTTTGTTAGCAAAATGCAAGTTACGAATCCTCTACC 2151  
Db 541 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560  
QY 2152 AATCTCCAGCCCAAGATCTCATCTCCAGTCATTGGTGGAGCACAACCTGCTGTTCTGT 2211  
Db 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580  
QY 2212 ACAATAGTGGCCCTATCTTTTGGAGTTTGGATGATACAGCTTATGATAGTATGAT 2271  
Db 581 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 600  
QY 2272 TGGCTAGGTTCTAAATCAAAAATCAATGTCCTGAAATACAGTTAGGAGTAAGCCCCA 2331  
Db 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620  
QY 2332 GCTAATGCCCCATCAGATTGACTCTTAGGGAATGAGATGCCCTAAGTATGGCTATCAAGA 2391  
Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 640  
QY 2392 AGCTCGAGCTTGGTGGATCCTTAATACAGCAAAATATGTCCTTATCTCTGAAGCT 2451

Db 641 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660  
QY 2452 ACATGGACTAAACTGGGTATAATCTGGGCTCAGCGAGTAGCTTCTTTGGTTCCAAAT 2511  
Db 661 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680  
QY 2512 AGTTATGGGATCCATTTTAGATATACGATCTCGCATCTCAGCAATTCAGCAAGTGTG 2571  
Db 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700  
QY 2572 GATGGGCTCTTATTTGTCGAGGATATTGGGTTCTCGAGTTTCGAAATTTCTTCTATCAT 2631  
Db 701 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720  
QY 2632 GACCCTGATGCTTTAGGTCAGGATATCGGTATATTAGTGGGGTTATCTCTTAGGACA 2691  
Db 721 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740  
QY 2692 AACTCCTACTTTGATCATCGATGTTGGTCTAGCATTTACCGAGTATTTGGTAGATCT 2751  
Db 741 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760  
QY 2752 AAAGATTATGTAGTGTGCTGTTCCAAATCATCATGCTTCATAGGATCCGTTTATCTATCT 2811  
Db 761 LysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyrLeuSer 780  
QY 2812 ACCCAACAGCTTTATGTGGATCCTTATTTGTTGGAGATGCGTTTATCCGTGCTAGTAC 2871  
Db 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800  
QY 2872 GGGTTGGGAATCAGCATATGAAAACCTCATATATACATTTGCGAGGAGCGATGTTGCT 2931  
Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820  
QY 2932 TGGGATAATACTGTCTGGCTGGAGATTCGAGCGGATACCGATTACCGATTGTGATTACTCA 2991  
Db 821 TrpAspAsnAsnCysLeuAlaGlyGlyIleGlyAlaGlyLeuProIleValIleThrPro 840  
QY 2992 TCTAAGCTCTATTGAAATGAGTTCGTCCTTTCGTGCAAGCTGAGTTCCTTATGCGCAT 3051  
Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860  
QY 3052 CATGAATCTTTTACAGAGAGGCGATCAAGCTCGGCGATTCCAGAGCGGACATCTCCTA 3111  
Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880  
QY 3112 AATCTATCAGTTCCTGTTGGAGTGAAGTTTCGATCGATGTTCTAGTACATCCTCTATAAA 3171  
Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900  
QY 3172 TATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231  
Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920  
QY 3232 ACGTCTCTATCCCATCAAGACATGAGACACAGATGCTTTCATTTTAGCAAGACATGGA 3291  
Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940  
QY 3292 GTTGTGGTTAGAGGATCTATGTTGCTTCTTAAACAAGTAATATAAGATATATGGCCAT 3351  
Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960  
QY 3352 GGAAGATGATGATTCGAGATGCTTCTCGAGCTATGTTTGAAGTCGAGGAAGTAGAGTC 3411  
Db 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980  
QY 3412 CGGTTC 3417  
Db 981 \*\*\*Phe 982

RESULT 13

US-09-598-419-176

; Sequence 176, Application US/09598419



Patent No. 6565856  
 GENERAL INFORMATION:  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Scholler, John  
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
 TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
 FILE REFERENCE: 210121.469C6  
 CURRENT APPLICATION NUMBER: US/09/598,419  
 CURRENT FILING DATE: 2000-06-20  
 NUMBER OF SEQ ID NOS: 357  
 SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 SEQ ID NO 176  
 LENGTH: 982  
 TYPE: PRT  
 ORGANISM: Chlamydia  
 FEATURE:  
 NAME/KEY: VARIANT  
 LOCATION: (1)...(982)  
 OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-598-419-176

Alignment Scores:  
 Pred. No.: 0 Length: 982  
 Score: 5084.00 Matches: 978  
 Percent Similarity: 99.8% Conservative: 2  
 Best Local Similarity: 99.6% Mismatches: 2  
 Query Match: 64.5% Indels: 0  
 DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-598-419-176 (1-982)

QY 472 ATGGTTCCTCAAGGATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATACGTGTT 531  
 DB 1 MetIleProGlnGlyIleTyAspGlyGluThrLeuThrValSerPheProTyThrVal 20  
 QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTCTCGCAGGAGAGTTAACTTAAATAATCTT 591  
 DB 21 IleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40  
 QY 592 GACAAATCTATTCAGCTTTGCTTTAAGTTGTTGGGAACCTTATAGGAGTTTACT 651  
 DB 41 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60  
 QY 652 GTTTTAGGAGAGACACTGTTGACTTTCGAGAACATAGGACTTCTCAAAATGGGCA 711  
 DB 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80  
 QY 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTTATTCAGGCTTTTAAAGAAATATCC 771  
 DB 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGlyPheLysGluLeuSer 100  
 QY 772 TTTTCCAAATTCATTCATTACTTTCGCGTACTGCTGCTCAACGACTAATAAGGGTAGC 831  
 DB 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 120  
 QY 832 CAGACTCCGAGCAGCACTACACGCTCTAATGGTACTATTTATTTCTTAAACAGATCTT 891  
 DB 121 GlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLysThrAspLeu 140  
 QY 892 TTGTTACTCAATATAGAGATGTTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGA 951  
 DB 141 LeuLeuLeuAsnAsnGlnLysPheSerPheTySerAsnLeuValSerGlyAspGlyGly 160  
 QY 952 GCTATAGATGCTAAGAGCTTAAACGGTTCAAGGAATTAGCAAGCTTTGTGTTCCCAAGAA 1011  
 DB 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180  
 QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTGTCAGTAGTCCAGTGTCTCTGCTATGCT 1071  
 DB 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200  
 QY 1072 AACGAGGCTCTATTGCTTTGTAGCGAATGTTGACGAGGTAAAGGGGGAGGATGCT 1131

DB 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 220  
 QY 1132 GCTGTTTCCAGGATGGCAGCAGGAGTGTTCATCTACTTCAACAGAGATCCAGTAGTA 1191  
 DB 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal 240  
 QY 1192 AGTTTTCAGAAATPACTGCGGTAGATTGATGGGAACGTAGCCGAGTAGGAGAGGG 1251  
 DB 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 260  
 QY 1252 ATTTACTCTACGGGAAGTTCCTTCTGATATATGGAACCTTGTCTTCTCAACAT 1311  
 DB 261 IleTySerTySerGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsn 280  
 QY 1312 GTTGCTTCTCTGTTTACATTCCTAAGCAACCAACAAAGTGGACAGGCTTCTAATACG 1371  
 DB 281 ValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300  
 QY 1372 AGTAATAATTACGGAGATGGAGAGCTATCTTCTGTAAGAATGGTGGCAAGCAGGATCC 1431  
 DB 301 SerAsnAsnTyGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320  
 QY 1432 AATAACTCTGGATCAGTTTCTTTCATCGAGAGGAGTAGTATTTCTTTAGTACGATGTA 1491  
 DB 321 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 340  
 QY 1492 GCTGCTGGGAAAGGGGAGCTATTATATGCAAAAAGCTCTCGGTGCTAACTGTGCCCCCT 1551  
 DB 341 AlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCysGlyPro 360  
 QY 1552 GTACAAATTTTAAAGGAATATCGCTAATGATGTGGAGCGAATTTATTTAGAGAACTCTGGA 1611  
 DB 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyLeuGlyGluSerGly 380  
 QY 1612 GAGCTCAGTTTCTGCTCATTTATGAGATATATTTTCGATGGGATCTTAAAGAACACA 1671  
 DB 381 GluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400  
 QY 1672 GCCAAGAGAACTGCTCCGATGTTAATGCGTAACGTCTCTCACAAGCCATTTCCGATG 1731  
 DB 401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420  
 QY 1732 GGATCCGGAGGGAATAACGACATTAAAGAGCTAAAGCAGGCGCATCAGATTCCTTTAAT 1791  
 DB 421 GlySerGlyGlyIleThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440  
 QY 1792 GATCCCATCAGATGGCAACCGAAATAACGACCGCCAGCCAGCTCTTCCAACTCTAAA 1851  
 DB 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys 460  
 QY 1852 ATTTACGATGTCAGGATACACAGGGATATTTGTTTGTCTAATGGAAGCAGTACTTTG 1911  
 DB 461 IleAsnAspGlyGluGlyTyThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480  
 QY 1912 TACCAAAATGTTACGATAGCAGCAAGGAAGGATTTGTTCTTCGTGAAAAGGCAAAATATCA 1971  
 DB 481 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGlyLysAlaLysLeuSer 500  
 QY 1972 GTCAATTCCTAAGTCAGACAGGAGTCTGTATATGGAAGCTGGAGTACATGGGAT 2031  
 DB 501 ValAsnSerLeuSerGlnThrGlySerLeuTyMetGluAlaGlySerThrLeuAsp 520  
 QY 2032 TTTGTAACCTCCACCAACCAACAGCTCTCCGCGCTAATCATGTTGATCAGCTTTC 2091  
 DB 521 PheValThrProGlnProGlnProAlaAlaAsnGlnIleThrLeuSer 540  
 QY 2092 AATCTGCAATTTGCTCTTCTTCTTGTAGCAAAACAAATCAGTTACGAAATCTCTTACC 2151  
 DB 541 AsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560  
 QY 2152 AATCTCTCCAGCGCAAGATCTCATCTCGAGTCATGTTGTTAGCAACATGCTGTTGTT 2211  
 DB 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580

```
QY 2212 ACAATTAGTGGCCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGTATGAT 2271
Db 581 ThrIleSerGlyProIlePheGluAspLeuAspThrAlaTyrAspArgTyrAsp 600
QY 2272 TGGCTAGGCTTCTAATCAAAAATCAATCTCCTGAAATTTACAGTTAGGAGCTAAAGCCCA 2331
Db 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620
QY 2332 GCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCCCTAAGTATGGCTATCAAGGA 2391
Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyThrGlnGly 640
QY 2392 AGCTGGAAGCTTGGCTGGATCTTAATACAGAAATAATGGTCTTATCTGAAAGCT 2451
Db 641 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660
QY 2452 ACATGGACTAAACTGGGTATATCTGGCTGAGCGTAGCTCTTTGGTTCCCAAT 2511
Db 661 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680
QY 2512 AGTTTATGGGATCCATTTAGATATACGATCTGGCATTGAGCAATTCAGCAAGTGTG 2571
Db 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700
QY 2572 GATGGCGCTCTTATTTGTCGAGGATATGGGTTCTGAGTTTCGAATTTCTTCTATCAT 2631
Db 701 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPheTyrHis 720
QY 2632 GACCGCGATCTTTAGCTCAGGATATCGGTATATTAGTGGGGTTATTCCTTAGGACA 2691
Db 721 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740
QY 2692 AACTCTCTACTTTGGATCATCGATGTTTGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751
Db 741 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
QY 2752 AAAGATTATGATGTGTGTTCCCAATCATCATGCTTGATAGGATCCGTTTATCTATCT 2811
Db 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780
QY 2812 ACCCAACAGCTTTATGTGGATCCTATTTGTTGCGAGATGCGTTTATCGGTAGCTAC 2871
Db 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800
QY 2872 GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGAGAGCGCATGTCGT 2931
Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluSerAspValArg 820
QY 2932 TGGGATAATACTGTCTGGCTGGAGATTTGAGCGGGATTAACGATTGTGATTACTCA 2991
Db 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
QY 2992 TCTAAGCTCTATTGATGATGTTGGTCTTTTCGTCAGCTGAGTTTCTTATCCGAT 3051
Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860
QY 3052 CATGAATCTTTTACAGAGAAAGCGCATCAAGCTCGGGCATTCAGAGCGGACATCTCCTA 3111
Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
QY 3112 AATCTATCAGTTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCTAATAA 3171
Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
QY 3172 TATAGCTTTATCGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTTACTGAGACA 3231
Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920
QY 3232 ACGCTCTATCCATCAAGAGACATGCAACAGATGCTCTTCATTTAGCAAGATGA 3291
Db 921 ThrLeuLeuSerHisGlnGlnThrTrpThrAspAlaPheHisLeuAlaArgHisGly 940
```

```
QY 3292 GTTCTGCTAGAGCATCTATGCTATCTCTTAACAAGTAATATAGAGTATATGCGCAT 3351
Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960
QY 3352 GGAAGATATGATGATCGAGATGCTTCTCGAGCTTATGGTGGTTCGAGGAGTAGAGTC 3411
Db 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980
QY 3412 CGGTTC 3417
Db 981 ***Phe 982

RESULT 14
US-10-197-220-169
; Sequence 169, Application US/10197220
; Patent No. 6919187
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Guderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/197,220
; CURRENT FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 169
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Chlamydia
US-10-197-220-169

Alignment Scores:
Pred. No.: 1,41e-300 Length: 670
Score: 3336.50 Matches: 651
Percent Similarity: 98.8% Conservative: 11
Best Local Similarity: 97.2% Mismatches: 7
Query Match: 42.3% Indels: 1
DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-197-220-169 (1-670)
QY 463 GCAGAAATCATGTTCTCTCAAGGAATTTACATGGGGAGAGTTTAACTGTATCATTTCCC 522
Db 1 AlaGluIleMetIleProGlnGlyTyrAspGlyGluThrLeuThrValSerPhePro 20
QY 523 TATCTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTCTCGCAGGAGAGTTAAACATTA 582
Db 21 TyrThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeu 40
QY 583 AAAATCTTGACAAATTTCTATTCAGCTTTGCCCTTTAAGTTGTTTTGGGAACCTTATTAGG 642
Db 41 LysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGly 60
QY 643 AGTTTATCTGTTTATAGGAGAGGACATCGTTGACTTTTCGAGAACATACGAGACTTCTACA 702
Db 61 SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr 80
QY 703 AATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTAGGGGTTTTTAAA 762
Db 81 AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys 100
QY 763 GAAATATCTCTTTTCCAAATTTGCAATTCATTTACTTCCGCTAGCTGCTGCTGCAACGACTAAT 822
Db 101 GluLeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsn 120
QY 823 AAGGTAGCCGAGACTCCCGACGACCACTACACCGTCTAATGGTACTACTATTTATTCTAAA 882
Db 121 AsnGlySerGlnThrProThrThrThrSerThrProSerAsnGlyThrThrIleThrSerLys 140
QY 883 ACAGATCTTTTCTTACTCAATAATCAGAAAGTTCTCATTTCTATAGTAATTTAGTCTCTGGA 942
```



```
QY 586 AATCTTGACAAATCTTAATTCGAGCTTTCCTTAAGTTGTTTGGAACTTATTAGGAGT 645
Db 41 AsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuGlySer 60
QY 646 TTTTACTGTTTGGGAGAGGACACTCGTTTGTACTTTCGAGAAACATACGACATCTTACAAAT 705
Db 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
QY 706 GGGGAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGAGGTTTAAAGAA 765
Db 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheGlyGlu 100
QY 766 TTATCCCTTTTCCAAATTCATTAATCTTACTTGGCTGCTGCTGCTGCAACGACTAATTAAG 825
Db 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrAsnLys 120
QY 826 GGTAGCCAGACTCCGACGACAACTCTACACCGTCTAATGGTACTATTATTCTTAAACA 885
Db 121 GlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLysThr 140
QY 886 GATCTTTTGTACTCAATAATAGAGAGTTCTCATTTCTATAGTAATTTAGTCTCTGGAGAT 945
Db 141 AspLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAsp 160
QY 946 GGGGAGCTATAGACTTAAGGCTTAAGGTTCAAGGATTAAGCAAGCTTGTGCTCTC 1005
Db 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
QY 1006 CAAGAAATACTGCTCAAGCTGATGGGAGCTCTCAAGTAGTCACCAAGTTTCTCTGCT 1065
Db 181 GlnGluAsnThrAlaGlnAlaAspGlyAlaCysGlnValAlaThrSerPheSerAla 200
QY 1066 ATGGCTAACGAGCTCTATTGCTTTCTAGCGAATGTTGCGAGGAGTAAGAGGGGAGG 1125
Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
QY 1126 ATTGCTGCTGTTACAGATGGGAGGAGGAGTGTCATCTACTTCAACAGAAATCGA 1185
Db 221 IleAlaAlaValGlnAspGlyGlnGlyValSerSerSerThrSerThrGluAspPro 240
QY 1186 GTAGTAAGTTTTCAGAAATACTCGGTAGAGTTGTGGAAGCTAGCCGAGTAGGA 1245
Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
QY 1246 GGAGGGATTTACTCTACGGGAACGTGCTTTCTGAAATAATGGAATAACCTTTCTCTC 1305
Db 261 GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280
QY 1306 AACAAATGTTGCTTCTCTGTTTACATTTGCTGCTAAGCAACCAAGTGGACAGGCTTCT 1365
Db 281 AsnAsnValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
QY 1366 AATACGAGTAATAATTCAGGAGATGGAGGAGCTATCTTCTGTAAGATGGTCGCAAGCA 1425
Db 301 AsnThrSerAsnAsnTyrGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320
QY 1426 GGATCCAAATAACTCGGATCAGTTTCTTTGATGGAGGGAGTAGTTTCTTTTAGTAGC 1485
Db 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValAlaPhePheSerSer 340
QY 1486 AATGTAGTCTGGGAAGGGGAGCTATTATTATGCCAAAAGCTCTCGTTCCTAACTGT 1545
Db 341 AsnValAlaAlaGlyLysGlyAlaIleTyrAlaLysLeuSerValAlaAsnCys 360
QY 1546 GSCCCTGTACAAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGATTTATTAGAGAA 1605
Db 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380
QY 1606 TCTGGAGAGCTCAGTTTATCTGCTGATTAATGGAGATATTATTTCGATGGGAATCTTAA 1665
Db 381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIlePheAspGlyAsnLeuLys 400
QY 1666 AGAACGCCAAAGAGATCTGCCGATGTTAATGGCGTAACGTGTCTCACAAGCCATT 1725
```

```
Db 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420
QY 1726 TCGATGGGATCGGAGGAGAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTC 1785
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
QY 1786 TTTAATGATCCATCGAGATGGCAACGGAATAATCAACGCCAGCGCAGTCTTCCAAACTT 1845
Db 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460
QY 1846 CTAAAAATTAAACGATGGTGAAGGATACACAGGGGATATTGTTTGTCTTAATGGAGCAGT 1905
Db 461 LeuLysIleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480
QY 1906 ACTTTGTACCAAAATGTTTACGATAGACCAAGGAAGGATGTTTCTTCGTGAAAAGGCAAAA 1965
Db 481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500
QY 1966 TTATCAGTGAATTC 1980
Db 501 LeuSerValAsnSer 505

RESULT 16
US-09-542-520-17
; Sequence 17, Application US/09542520
; Patent No. 6887843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John L.
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; PRIOR FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-542-520-17

Alignment Scores:
Pred. No.: 2,35e-227 Length: 505
Score: 2547.00 Matches: 504
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 32.3% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-542-520-17 (1-505)
QY 466 GAAATCATGGTTCCTCAAGGATTTACGATGGGAGAGTTAACTGTATCATTTCCCTAT 525
Db 1 GluIleMetValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20
QY 526 ACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTCTGCGAGGAGAGTTAACTTAAAA 585
Db 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
QY 586 AATCTTGACAAATCTTAATTCGAGCTTTCCTTTAAGTTGTTTGGAACTTATTAGGAGT 645
Db 41 AsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuGlySer 60
QY 646 TTTTACTGTTTGGGAGAGGACACTCGTTGACTTTCGAGAAACATACGACATCTTCAAAAT 705
Db 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
QY 706 GGGGAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGAGGTTTAAAGAA 765
Db 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheGlyGlu 100
```

QY 766 TTATCCTTTTCCAAATTCGAATTCATTACTTCCGCTACTCCCTGCTGCAACGACTAATAAG 825  
Db 101 LeuSerPheSerAsnGlySerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120  
QY 826 GGTAGCCAGACTCCGACGACAACTCTACACCGTCTAATGGTACTATTATTCTTAAACA 885  
Db 121 GlySerGlnThrProThrThrThrSerThrProSerAsnGlyThrIleThrSerLysThr 140  
QY 886 GATCTTTTGTACTCAATAATGAGAGAGTTCTCATTTCTATAGTAATTTAGTCTCTCGAGAT 945  
Db 141 AspLeuLeuLeuLeuAsnAsnGlnLysPheSerPheThrSerAsnLeuValSerGlyAsp 160  
QY 946 GGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGCAATTAGCAAGCTTTGTCTCTTC 1005  
Db 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180  
QY 1006 CAAGAAATACTGCTCAAGCTGATCGGGAGCTTCTCAAGTAGTCAACAGTTTCTCTGCT 1065  
Db 181 GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAla 200  
QY 1066 ATGGCTACGAGGCTCTATTGCTCTTTGTAGCAATGTTGCAGGAGTAGAGGGGGAGG 1125  
Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220  
QY 1126 ATTGCTGCTGTTACGATGGGCGAGGAGTGTCTATCTACTTCAACAGAGTCA 1185  
Db 221 IleAlaAlaValGlnAspGlyGlnGlyValSerSerSerThrThrGluAspPro 240  
QY 1186 GTAGTAAGTTTTCCAGAAATACTCGGCTAGAGTTTGTATGGGAAGTACCGAGTAGGA 1245  
Db 241 ValValSerPheSerArgAsnThrAlaValAlaGluPheAspGlyAsnValAlaArgValGly 260  
QY 1246 GGAGGATTTACTCTACGGGAACGTTGCTTTCTCTGTAATATGGAATAACCTTTGTTCTC 1305  
Db 261 GlyGlyIleThrSerThrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280  
QY 1306 AACATGTTGCTTCTCTGTTTACATTTGCTGTAGCAACCAACAGTGGACAGCTTCT 1365  
Db 281 AsnAsnValAlaSerProValThrIleAlaAlaLysGlnProThrSerThrGlnAlaSer 300  
QY 1366 AATACGAGTAATAATTACGAGATGGAGAGCTATCTTCTGTAAGAATGGTCGCAAGCA 1425  
Db 301 AsnThrSerAsnAsnThrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320  
QY 1426 GAATCCAATACTCTGGATCAGTTTCTTTGTATGGAGAGGAGTAGTTTCTTTAGTAGC 1485  
Db 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340  
QY 1486 AATGTAGCTGCTGGGAAGGGAGCTATTTATGCCAAAAGCTCTCGGTTCTTAACCTGT 1545  
Db 341 AsnValAlaAlaGlyLysGlyGlyAlaIleThrAlaLysLysLeuSerValAlaAsnCys 360  
QY 1546 GGCCTGTACAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGATTTATTAGAGAA 1605  
Db 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleThrLeuGlyGlu 380  
QY 1606 TCTGGAGAGCTCAGTTTATCTGCTGATATGAGATATTTATTTCAATGGGAATCTTAAA 1665  
Db 381 SerGlyGluLeuSerLeuSerAlaAspThrGlyAspIlePheAspGlyAsnLeuLys 400  
QY 1666 AGAACGCAAGAGAGATCTCCGATGTTAATGGGTAACTGCTGCTCCTCACAGCCATT 1725  
Db 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420  
QY 1726 TCGATGGGATCGGAGCGGAAATAACGACATTAAGAGCTAAAGCAGGCGATCAGATTCTC 1785  
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440  
QY 1786 TTTAATGATCCCATCGAGATGGCAACCGGAATAACCGACCGCAGCTTCTCCAACTT 1845  
Db 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460

QY 1846 CTAATAATTAACGATGGTGAAGGATACACAGGGGATATTTGTTTGTCTAATGGAGCAGT 1905  
Db 461 LeuLysIleAsnAspGlyGluGlyThrThrGlyAspIleValPheAlaAsnGlySerSer 480  
QY 1906 ACTTTGTACCAAAATGTTACGATAGACGACGAGGAGTGTCTTCTGTAAGGCAAAA 1965  
Db 481 ThrLeuThrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgLysAlaLys 500  
QY 1966 TTATCAGTGAATTTCT 1980  
Db 501 LeuSerValAspSer 505  
RESULT 17  
US-09-612-402B-36  
; Sequence 36, Application US/09612402B  
; Patent No. 6642023  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; Patent No. 6642023  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/09/612,402B  
; CURRENT FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 36  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-09-612-402B-36  
Alignment Scores:  
Pred. No.: 4,32e-209 Length: 458  
Score: 2350.00 Matches: 458  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 29.8% Indels: 0  
DB: 2 Gaps: 0  
US-10-701-844-1 (1-4435) x US-09-612-402B-36 (1-458)  
QY 1030 GGGGAGCTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCTAACGAGGCTCTATTGCC 1089  
Db 1 GlyGlyAlaCysGlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAla 20  
QY 1090 TTTGTAGCGAATGTTTCAGGAGTAGAGGGGAGGAGTTGCTGCTTTCAGGATGGGAG 1149  
Db 21 PheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGln 40  
QY 1150 CAGGAGGTGTCATCTACTTCAACAGAGTCCAGTAGTAGTTTTCAGAGTAATACT 1209  
Db 41 GlnGlyValSerSerThrSerThrGluAspProValValSerPheSerArgAsnThr 60  
QY 1210 GGGGTAGAGTTTGTAGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTTACGGGAAC 1269  
Db 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyIleThrSerThrGlyAsn 80  
QY 1270 GTTGTCTTCTCAATAATGGAAAAACCTTTGTTTCTCAACAATGTTGCTTCTCTCTTTAC 1329  
Db 81 ValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValThr 100  
QY 1330 ATTGCTGTAGCAACCAACAGTGGAGCGGCTTCTTAATACAGTAATTAATTCAGGAT 1389  
Db 101 IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnThrGlyAsp 120  
QY 1390 GGAGGAGCTATCTTCTGTAAGAAATGGTCGCAAGCAGGATCCAATAACTCTCGATCAGTT 1449  
Db 121 GlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal 140  
QY 1450 TCCTTTGATGGAGGAGGAGTAGTTTCTTTTAGTCAATGTAGCTGCTCGGAAAGGGGA 1509

```
141 SerPheAspGlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGly 160
1510 GCTATTATGCCCCAAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAAATTTTAAAGGAAT 1569
161 AlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn 180
1570 ATCGCTAATCATGTGAGCGGATTTATTAGGAGATCTCGAGAGCTCAGTTTATCTGCT 1629
181 IleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAla 200
1630 GATTATGAGATATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCC 1689
201 AspTyrGlyAspIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAla 220
1690 GATGTTAATCGGTAACGTGCTCTCAAGCCATTTTCGATGGATCGGAGGGGAATA 1749
221 AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIle 240
1750 ACGACATTAAGCTAAAGCAGGCGCATCAGATTCTCTTAAATGATCCCATCGAGATGCCA 1809
241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAla 260
1810 AACGGAATAACACGACGAGCGAGCTCTTCCAACTCTTAAATTAACGATGGTGAAGGA 1869
261 AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGly 280
1870 TACACAGGGGATATGTTTGTCTAATGAAGAGAGTACTTTGTACCAAAATGTTACGATA 1929
281 TyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIle 300
1930 GAGCAAGGAAGGATGTTCTTCGTAAGAACCAAAATTTATCAGTGAATTTCTTAAGTCAG 1989
301 GluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGln 320
1990 ACAGTGGAGCTGTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACCA 2049
321 ThrGlySerLeuTyrMetGluAlaGlySerThrTrpAspPheValThrProGlnPro 340
2050 CCACACAGCTCTCGCGCTAATCAGTTGATCAGGCTTTCATCTGCATTTGCTCTTT 2109
341 ProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu 360
2110 TCTCTTTGTAGCAACAAGTCAGTACGAATCTCTCTACCAATCTCTCCAGCGCAAGAT 2169
361 SerSerLeuLeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAsp 380
2170 TCTCATCTGCAGTCATTGGTAGCAACTGCTGCTTCTGTACAAATTAGTGGGCTATC 2229
381 SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIle 400
2230 TTTTGTAGGATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAA 2289
401 PhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGln 420
2290 AAAATCANTGCTCGAATACAGTTAGGACTAAGCCCGAGCTAAGCCCATCGATCAGAT 2349
421 LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAsp 440
2350 TTGACTTAGGAAATGAGATGCTAAGTATGCTATCAAGGAAGCTGGAAAGCTT 2403
441 LeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeu 458
```

```
RESULT 18
US-09-542-520-36
; Sequence 36, Application US/09542520
; Patent No. 6887843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John L.
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
```

```
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; PRIOR FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-542-520-36

Alignment Scores: 4.32e-209 Length: 458
Pred. No.: 2350.00 Matches: 458
Score: 2350.00 Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 29.8% Gaps: 0
DB: 2

US-10-701-844-1 (1-4435) x US-09-542-520-36 (1-458)

QY 1030 GGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGCTAACGAGGCTCTTATTGCC 1089
Db 1 GlyGlyAlaCysGlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAla 20
QY 1090 TTTGTAGCGAATGTTGAGGAGTAAGAGGGGGAGGATTCCTGCTGTTCAGGATGGGAG 1149
Db 21 PheValAlaAsnValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGln 40
QY 1150 CAGGAGGTGCATCATCTTCTCAACAGAGATCCAGTAGTAACTTTTCCAGAAATCT 1209
Db 41 GlnGlyValSerSerThrThrGluAspProValValSerPheSerArgAsnThr 60
QY 1210 GCGGTAGAGTTGATGGAGACGTAGCCGAGTAGGAGGAGGATTTTACTCTACGGGAAC 1269
Db 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsn 80
QY 1270 GTTGCTTTCTCGAATAATGGAAAAACCTTGTCTCAACAATGTTGCTTCTCTGTTTAC 1329
Db 81 ValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyr 100
QY 1330 ATTGCTGTAAGCAACCAACAGTGGAGAGGCTTCTTAACGAGTAAATTAATGAGGAT 1389
Db 101 IleAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAsp 120
QY 1390 GGAGGAGTATCTTCTGAATGTGAATGGCGCAAGCAGGATCCAATACTCTGGATCAGTT 1449
Db 121 GlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal 140
QY 1450 TCCTTTGATGAGAGGAGTAGTTTCTTTTAGTAGCAATGATGCTGTGGAAAGGGGGA 1509
Db 141 SerPheAspGlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGly 160
QY 1510 GCTATTATGCAAAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAAATTTTAAAGGAAT 1569
Db 161 AlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn 180
QY 1570 ATCGCTAATCATGTGAGCGGATTTATTAGGAGATCTCGAGAGCTCAGTTTATCTGCT 1629
Db 181 IleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAla 200
QY 1630 GATTATGAGATATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCC 1689
Db 201 AspTyrGlyAspIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAla 220
QY 1690 GATGTTAATGCGTAACTGCTCTCAAGCCATTTTCGATGGATCGGAGGGGAATA 1749
Db 221 AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIle 240
QY 1750 ACGACATTAAGCTAAAGCAGGCGCATCAGATTCTCTTAAATGATCCCATCGAGATGCCA 1809
Db 241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAla 260
```





```
/ APPLICANT: Pace, John L.
/ TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
/ FILE REFERENCE: 7969-076-999
/ CURRENT APPLICATION NUMBER: US/09/542,520
/ CURRENT FILING DATE: 2000-04-03
/ PRIOR APPLICATION NUMBER: PCT/US98/20737
/ PRIOR FILING DATE: 1998-10-01
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 325
/ TYPE: PRT
/ ORGANISM: Chlamydia
/ US-09-542-520-37

Alignment Scores:
Pred. No.: 4, 4e-152 Length: 325
Score: 1735.00 Matches: 325
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.0% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-542-520-37 (1-325)
QY 2443 CTGAAGCTACATGACATAAAGCTGGGTATATCTCGGCTGAGGAGTAGCTCTTTG 2502
Db 1 LeuYsAlaThrTrpThrylsThrGlyTyraAnProGlyProGluArgValAlaSerLeu 20
QY 2503 GTTCCAAATAGTTTGGGATCCATTTAGATATACGATCGCCATTCAGCAATCAA 2562
Db 21 ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln 40
QY 2563 GCAAGTGTGGATGGCGCTCTTATTGTCGAGGATATGCGTTCCTGAGTTTCGAATTC 2622
Db 41 AlaSerValAspGlyArgSerTyrcysArgGlyLeuTrpValSerGlyValSerAnPhe 60
QY 2623 TTCTATCATGACCGGATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATCC 2682
Db 61 PheTyrcysAspArgAspAlaLeuGlyGlnGlyTyrcysArgGlyLeuTrpValSerGlyTyrcys 80
QY 2683 TTAGGAGCAATCTCTCTGATCATCGATGTTGGTCTAGCATTTACCAAGTATT 2742
Db 81 LeuGlyAlaAsnSerTyrcysGlySerMetPheGlyLeuAlaPheThrGluValPhe 100
QY 2743 GGTAGATCTAAAGATTATGATGTCGTTCCCAATCATCATGCTTCATGATGATCCGTT 2802
Db 101 GlyArgSerLeuAspTyrcysArgGlySerAsnHisAlaCysIleGlySerVal 120
QY 2803 TATCTATCTACCAACAAAGCTTTATGTCGATCTTATTGTCGGAGATCGGTTTATCCGT 2862
Db 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrcysArgGlyLeuPheGlyAspAlaPheIleArg 140
QY 2863 GCTAGCTAGGCTGGGATCAGCATATGAACTCATATACATATGCTTCAGGAGAGC 2922
Db 141 AlaSerTyrcysGlyPheGlyAsnGlnHisMetTyrcysSerTyrcysPheAlaGluGluSer 160
QY 2923 GATGTTTGGGATTAATCTGTCGCTGAGAGATGAGCGGATTTACCGATTGTG 2982
Db 161 AspValArgTrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleVal 180
QY 2983 ATTACTCCATCTAAGCTCTATTGAATGATGTCGTCCTTTCTGTCGAGCTAGTTTCT 3042
Db 181 IleThrProSerTyrcysLeuTyrcysGlySerLeuAsnGluLeuArgProPheValGlnAlaGluPheSer 200
QY 3043 TATCGGATCATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCATTTCAAGAGCGGA 3102
Db 201 TyrAlaAspPheIleGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheTyrcysGly 220
QY 3103 CATCTCTAAATCTATCATGTTCTGTTGGAGTGAAGTTTGTGATCGATGTTCTAGTACACAT 3162
Db 221 HisLeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHis 240

/ APPLICANT: Steven P. Fling
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
/ TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
/ FILE REFERENCE: 210121.469C7
/ CURRENT APPLICATION NUMBER: US/09/620,412C
/ CURRENT FILING DATE: 2000-07-20
/ NUMBER OF SEQ ID NOS: 363
/ SOFTWARE: FastSeq for Windows Version 3.0/4.0
/ SEQ ID NO 325
/ LENGTH: 631
/ TYPE: PRT
/ ORGANISM: Chlamydia trachomatis
/ US-09-620-412C-325

Alignment Scores:
Pred. No.: 1, 68e-138 Length: 631
Score: 1590.50 Matches: 316
Percent Similarity: 88.7% Conservative: 7
Best Local Similarity: 86.8% Mismatches: 19
Query Match: 20.2% Indels: 22
DB: 2 Gaps: 4

US-10-701-844-1 (1-4435) x US-09-620-412C-325 (1-631)
QY 3391 TTGAGTCGAGGAGTAGAGTCGGTCTTAAATAATAT-----TGGTAGATAGTTAA 3441
Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrcysArgTyrcysArgPro----- 146
QY 3442 GTGTTAGCGATGCGCTTTTCTTTGAGATCTACATCATTTGTTTGTAGCTTTGTCT 3501
Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164
QY 3502 TCCTATTGCTATGGATTGCGAGCTCTCTCAAGTGTTAACGCTTAATGTAACTACTCT 3561
Db 165 SerTyrcysTyrcysGlyPheAlaSerSerProGlnValLeuThrProAsnValThrPro 184
QY 3562 TTTAAGGAGGAGCATGTTTACTTGAATGAGATCGCTGCTTTGTCAATGTCTATCGAGA 3621
Db 185 PheTyrcysGlyAspAspValTyrcysLeuAsnGlyAspCysAlaPheValAsnValTyrcys 204
QY 3622 GCTGAAGAGGTTTCCATTATCTCAGCTAATGCGCAATTTAAACGATTACCGGACAAAAC 3681
Db 205 AlaGluAsnGlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 224
QY 3682 CATACATTATCATTTTACAGATTTCTCAAGGCGCATGTTCTTCAAAATATGCTTCAATTC 3741
Db 225 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrcysAlaPheIleSer 244
```

QY 3742 GCAGGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCT 3801  
 DB 245 AlaGlyGluThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSer 264  
 QY 3802 TCGGAGAAAGGGAATGATCTCCGGGAAACCGTGAGTATTTCGAGCAGCGAAGTG 3861  
 DB 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284  
 QY 3862 ATTTCCTGGGATCACTCCGTTGGGTTATCTCTTATCTACTGTGTCACCTCATCATCA 3921  
 DB 285 IlePheTrpAspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304  
 QY 3922 ACTCCGCTGCT-----CCACAGTTAGT 3945  
 DB 305 ThrProProAlaProAlaProAlaProAlaAAserSerSerLeuSerProThrValSer 324  
 QY 3946 GATGCTCGGAAAGGCTATTTCCTGTAGAGACTAGTTTGGAGATCTCAGCGTCACAA 4005  
 DB 325 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 344  
 QY 4006 AAGGGGTCATGTCGATTAATATGCGGGAATTCGGAACAGTTTTCGAGTAAAGAT 4065  
 DB 345 LysGlyValMetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSer 364  
 QY 4066 AATAAATGCTGTCGATTAATATGCGGGAATTCGGAACAGTTTTCGAGTAAAGAT 4125  
 DB 365 AsnAsnAsnAlaGlySerGlyLysThrValSerIlePheGlyThrValPheArgGly 382  
 QY 4126 AGTAAATGCTGTCGATTAATATGCGGGAATTCGGAACAGTTTTCGAGTAAAGAT 4185  
 DB 382 rValLysAsnCysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyGly 402  
 QY 4186 AGTGGTTTATAAAGGCAATGTCGTTTTCAAAGCAATGAAGAGGCAATATTCCTCCAGG 4245  
 DB 402 yValValTyrLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgGly 422  
 QY 4246 GAACACAGCATACGATTAAGGATTCCTGCTACTACTACTACTACTACTACTACTACT 4305  
 DB 422 yAsnThrAlaTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGly 442  
 QY 4306 GACAGGAGCGGTGGAGGATTTATTCCTCTCAGAGATGATTCGTAAAGTTTGAAGCAA 4365  
 DB 442 uThrGlyGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAs 462  
 QY 4366 TAAAGTTCATTGTTTTCGATTAACATTTGCAAAAGGAGGAGCGGGAAGCATCTCAAC 4425  
 DB 462 nLysGlySerIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuTh 482  
 QY 4426 GAAAGATTC 4435  
 DB 482 rLysGluPhe 485

RESULT 22

; US-09-598-419-325  
 ; Sequence 325, Application US/09598419  
 ; Patent No. 6565856  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Scholler, John  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
 ; ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
 ; ; FILE REFERENCE: 210121.469C6  
 ; CURRENT APPLICATION NUMBER: US/09/598,419  
 ; CURRENT FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 357  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 325  
 ; LENGTH: 631  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia trachomatis  
 ; US-09-598-419-325

Alignment Scores:

Pred. No.: 1.68e-138 Length: 631  
 Score: 1590.50 Matches: 316  
 Percent Similarity: 88.7% Conservative: 7  
 Best Local Similarity: 86.8% Mismatches: 19  
 Query Match: 20.2% Indels: 22  
 DB: 2 Gaps: 4  
 US-10-701-844-1 (1-4435) x US-09-598-419-325 (1-631)  
 QY 3391 TTGAGTCCGAGGAGTAGAGTCCGGTTCATAAATAATAT-----TGGTTAGATAGTTAA 3441  
 DB 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisTrpArgPro----- 146  
 QY 3442 GTGTTAGCATGCTCTTTTGTAGATCTACATCATTTGTTTGTGTTTGTGTTGTTGTT 3501  
 DB 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164  
 QY 3502 TCTATTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3561  
 DB 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrPro 184  
 QY 3562 TTTAAGGAGAGCATGTTTACTTGAATGAGAGTCCGCTTTTGTCAATGCTATGTCAGGA 3621  
 DB 185 PheLysGlyAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGly 204  
 QY 3622 GCTGAAGAGGTTTCGATTAATCTCAGCTAATGCGCAATTTTAAACGATTACCGGACAAA 3681  
 DB 205 AlaGluAsnGlySerIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 224  
 QY 3682 CATACATTATCATTTACAGATCTCAAGGCCAGTTCTTCAAAATATGCTTCATTCTCA 3741  
 DB 225 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSer 244  
 QY 3742 GCAGGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCT 3801  
 DB 245 AlaGlyGluThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSer 264  
 QY 3802 TCGGAGAAAGGGAATGATCTCCGGGAAACCGTGAGTATTTCGAGCAGCGAAGTG 3861  
 DB 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284  
 QY 3862 ATTTCCTGGGATCACTCCGTTGGGTTATTCCTTATCTACTGTGTCACCTCATCATCA 3921  
 DB 285 IlePheTrpAspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304  
 QY 3922 ACTCCGCTGCT-----CCACAGTTAGT 3945  
 DB 305 ThrProProAlaProAlaProAlaProAlaAAserSerSerLeuSerProThrValSer 324  
 QY 3946 GATGCTCGGAAAGGCTATTTCCTGTAGAGACTAGTTTGGAGATCTCAGCGTCACAA 4005  
 DB 325 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 344  
 QY 4006 AAGGGGTCATGTCGATTAATATGCGGGAATTCGGAACAGTTTTCGAGTAAAGAT 4065  
 DB 345 LysGlyValMetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSer 364  
 QY 4066 AATAAATGCTGTCGATTAATATGCGGGAATTCGGAACAGTTTTCGAGTAAAGAT 4125  
 DB 365 AsnAsnAsnAlaGlySerGlyLysThrValSerIlePheGlyThrValPheArgGly 382  
 QY 4126 AGTAAATGCTGTCGATTAATATGCGGGAATTCGGAACAGTTTTCGAGTAAAGAT 4185  
 DB 382 rValLysAsnCysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyGly 402  
 QY 4186 AGTGGTTTATAAAGGCAATGTCGTTTTCAAAGCAATGAAGAGGCAATATTCCTCCAGG 4245  
 DB 402 yValValTyrLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgGly 422  
 QY 4246 GAACACAGCATACGATTAAGGATTCCTGCTACTACTACTACTACTACTACTACTACT 4305  
 DB 422 yAsnThrAlaTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGly 442

QY	4306	GACAGGAGCGGTGGAGGAGCTTATTTCCTCTCCAGATGATTCCTGTAAGTTTGAAGGCA	4365
Db	442	uThrGlyGlyGlyGlyValIleCysSerProAspSerValIysPheGluGlyAs	462
QY	4366	TAAAGGTCTTATTGTTTTTCATTACAACTTTGC AAAAGGAGGAGCGGAGCATCTCTAAC	4425
Db	462	nLysGlySerIleValPheAspTyrAenPheAlaLysGlyArgGlyGlySerIleLeuTh	482
QY	4426	GAAAGATTC	4435
Db	482	rLysGluPhe	485
RESULT 23			
US-10-197-220-95			
; Sequence 95, Application US/10197220			
; Patent No. 6919187			
; GENERAL INFORMATION:			
; APPLICANT: Bhatia, Ajay			
; APPLICANT: Guderian, Jeff			
; APPLICANT: Skeiky, Yasir A. W.			
; APPLICANT: Maisonneuve, Jean-Francois L.			
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT			
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION			
; FILE REFERENCE: 210121.515C3			
; CURRENT APPLICATION NUMBER: US/10/197,220			
; CURRENT FILING DATE: 2002-07-15			
; NUMBER OF SEQ ID NOS: 175			
; SEQ ID NO 95			
; LENGTH: 1016			
; TYPE: PRT			
; ORGANISM: Chlamydia trachomatis serovar D			
US-10-197-220-95			

```

3523 AGCTCTCTCAAGTGTAAAGCGCTAATGTAAACCACTCTTTTAAAGGAGACGATGTTTAC 3582
Db 1 SerSerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspValTyr 20

3583 TTGAATGGAGACTGCGCTTTTGTCAATGCTATGACGAGAGCTGAAGAAGTTCGATATTC 3642
Db 21 LeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIle 40

3643 TAGCTTAATGGGACAAATTAACGATACCGGACAAACCAATCATATCATATACAGAT 3702
Db 41 SerAlaAsnGlyAspAsnLeuThrThrThrGlnAsnHisThrLeuSerPheThrAsp 60

3703 TCTCAAGGCCAGCTCTCTCAAAATATGCTTCATTTACGACGAGACACATCTACTG 3762
Db 61 SerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeu 80

3763 AGAGATTTTTCGAGTCTGATGTTCTCGAAATATGTTCTTTCGCGAGAAAGGGAATGATC 3822
Db 81 LysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlyMetIle 100

3823 TCCGGGAAACCGTGAGTATTTCCGGAGCAGCGGAAGTGAATTTCTGGGATACCTCCGTG 3882
Db 101 SerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrAspAsnSerVal 120

3883 GGGTATCTCTTATCTACTGTGCGCACTCATCATCACTCCGCTGCT----- 3933
Db 121 GlyTyrSerProLeuSerIleValProAlaSerThrProThrProAlaProAlaPro 140

3934 -----CCAACAGTTAGTGTCTCGGAAGGCTTATT 3966
Db 141 AlaProAlaAlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIle 160

3967 TTTTCTGTAGACTAGTTGGAGATCTCAGCGCTCAAAAGGGGTCAATGTTTCGATAT 4026
Db 161 PheSerValGluThrSerLeuGluIleSerGlyValLysGlyValMetPheAspAsn 180

4027 AATCGCGGAATTTCCGACAGTTTTCGAGGTAAAGTAATTAATATGCTGCTGCTGGA 4086
Db 181 AsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsnAlaGlySerGly 200

4087 GGCAGTGGGTTCGCTACACCATCAAGTACGACTTTTACAGTTAAAAAAGTAAAGGAA 4146
Db 201 GlySerGlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGly 218

4147 AGTTTCTTTCCAGATAACCTAGCTCTTCGCGAGCGGAGTGGTGTATTAAGGCAATGT 4206
Db 218 sValSerPheThrAspAsnValAlaSerCysGlyGlyValValTyrLysGlyThrVa 238

4207 GCTTTTCAAGACAAATGAAGGAGCATATTTCTTCGAGGGAACACAGCATACGATGTTT 4266
Db 238 lLeuPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLe 258

4267 AAGGATTTCTGCTACTAATCAGGATCAGATAACGACGACGAGCGGCTGGAGAGT 4326
Db 258 uGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyGlyVa 278

4327 TATTGCTCTCCAGATGATTCTGTAAAGTTTGAAGCAATAAAGTTTCTATTGTTTGA 4386
Db 278 lIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySerIleValPheAs 298

4387 TTACAATTTGCAAAAGCGAGCGGAGCAAGCATCTTAACGAAGATTC 4435
Db 298 pTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPhe 314

```

RESULT 25

```

US-09-438-185A-455
; Sequence 455, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald

```

```

; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 455
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPN0453
US-09-438-185A-455

Alignment Scores:
Pred. No.: 1,37e-117 Length: 999
Score: 1366.50 Matches: 354
Percent Similarity: 49.7% Conservative: 177
Best Local Similarity: 33.1% Mismatches: 430
Query Match: 17.3% Indels: 107
DB: 2 Gaps: 28

US-10-701-844-1 (1-4435) x US-09-438-185A-455 (1-999)

```

```

QY 322 ATTAATAATTAATTTTATGAAGCGGAGTAATTAATTTTATCTCTC----- 369
Db 3 ValLeuLeuTyrLeuPhePheTyrSerLeuSerLeuIleCysArgIleIleTyrPheHis 22

QY 370 AGCTTTTGTGTGAACGCTCTTCCATAGTCTTCTTCTTCAATGATCTTAGCTTAT 429
Db 23 LeuTyrValGlnMetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAlaPro 42

QY 430 TCTTGTGCTCTTTAAATGCGGGGATATGCAGCAGAAATCATGTCTCTCAAGGAATT 489
Db 43 CysPheAlaSerThr-----AlaPheThrValGluValIleMetProSerGluAsn 59

QY 490 TAGCATGGGAGACGCTTAATCTGATCATTTCCCTATCTATTATAGAGATCCGAGTGGG 549
Db 60 PheAspGlySerSerGlyLysIle---PheProTyrThrThrLeuSerAspProArgGly 78

QY 550 ACTACTGTTTTTCTCGAGAGAGTTAACAATTAATAAATCTTGACAATCTTATTCGACT 609
Db 79 ThrLeuCysIlePheSerGlyAspLeuTyrIleAlaAsnLeuAspAsnAlaIleSerArg 98

QY 610 TTGCTTTTAAGTTGTTTGGGAACTTATTAGGAGTTTACTGTTTTCGAGGAGACAC 669
Db 99 ThrSerSerSerCysPheSerAsnArgAlaGlyAlaLeuGlnIleLeuGlyLysGly 118

QY 670 TCGTTGACTTTCGAGAACATACGACTTCTACAATATGGGCGAGCTCTAAGTAATAGCGCT 729
Db 119 ValPheSerPheLeuAsnIleArgSerSerAlaAspGlyAlaAlaIleSerSerValIle 138

QY 730 GCTGATGGA-----CTGTTTACTATTGAGGGTTTAAAGAAATATCTTTTCC 777
Db 139 ThrGlnAsnProGluLeuCysProLeuSerPheSerPheSerGlnMetIlePheAsp 158

QY 778 AATTGCAATTCATTACTTGGCGTACTGCTGCTGCAACGACTAATAAGGTCAGCAGACT 837
Db 159 AsnCysGluSerLeuThr-----SerAspThr 167

QY 838 CCGACGACACACTCTACACCGCTCTAATGGTACTATTATTCTAAAAACAGATCTTTTGT 897
Db 168 SerAlaSerAsnValIleProHisAlaSerAlaIleTyrAlaThrThrProMetLeuPhe 187

QY 898 CTCATAATGAGAAAGTTCTCATTCTATAGTAATTTAGTCTCTCGAGATGGGAGCTATA 957
Db 188 ThrAsnAsnAspSerIleLeuPheGlnTyrAsnArgSerAlaGlyPheGlyAlaAlaIle 207

```









; ORGANISM: Chlamydia pneumoniae			
US-09-428-122-2			
Alignment Scores:			
Pred. No.:	1,46e-95	Length:	928
Score:	1129.00	Matches:	312
Percent Similarity:	45.7%	Conservative:	171
Best Local Similarity:	29.5%	Mismatches:	401
Query Match:	14.3%	Indels:	174
DB:	2	Gaps:	30
US-10-701-844-1 (1-4435) x US-09-428-122-2 (1-928)			
QY	382	ATGCAACGCTCTTTCATAGTCTCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT	441
DB	1	MetLysSerSerPheProLysPheValPheSerThrPheAlaIlePheProLeuSerMet	20
QY	442	TTAAATGGGGGGATATGCAGAGAA---ATCATGGTTCTTCAAGGAATTTACGATGGG	498
DB	21	Ile-----AlaThrGluThrValLeuAspSerSerAlaSerPheAspGly	35
QY	499	GAGACGTTAACTGATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTT	558
DB	36	AsnLys---AsnGlyAsnPheSerValArgGluSerGlnGluAspAlaGlyThrThrTyr	54
QY	559	TTTTCTGAGGAGATTACATTAATAAATCTTGCAATCTTATTCGACGTTTGCCTTTA	618
DB	55	LeuPheLysGlyAsnValThrLeuGluAsnIleProGlyThrGlyThrAlaIleThrLys	74
QY	619	AGTTGTTTTGGGAACTTATTAGGAGTTTTTACTGTTTATAGGAGAGACACTCGTTGACT	678
DB	75	SerCysPheAsnAsnThrLysGlyAspLeuThrPheThrGlyAsnGlyAsnSerLeuLeu	94
QY	679	TTTCGAGAACATCGGACTTCTACA---AATGGGGCAGCTCTAAGTAATAGCGCTGCTGAT	735
DB	95	PheGlnThrValAspAlaGlyThrValAlaGlyAlaAlaValAsnSerSerValValasp	114
QY	736	GGACTGTTTACTATTGAGGGTTTAAAGATTTATCTCTTTCATTTCCATTTCAATTCTT	795
DB	115	LysSerThrThrPheIleGlyPheSerSerLeuSerPhe-----IleAla	129
QY	796	GGCGTACTCGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGCAACATCTACA	855
DB	130	SerProGlySerSerIleThrThrGlyLysGlyAlaValSerCysSerThrGlySer---	148
QY	856	CCGCTCTAATGGTACTATTATTCTAAACACAGACTTTTGTGTTACTCAATAATGAGAAGTTC	915
DB	149	-----LeuSerLeuThrLysAsnValSerLeu	157
QY	916	TCATTCATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGACTTAACG	975
DB	158	LeuPheSerLysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaLysThrLeuSer	177
QY	976	GTTCAAGGAATTAGCAAGCTTTGTGCTCCAGAAAATACTCTCAAGCTGATGGGGA	1035
DB	178	LeuThrGlyThrThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGlyGly	197
QY	1036	GCTTGTCAAGTAGTCACGACTTCTCTGCTATGGCTAAACGAGCTCCTATTTGCTTTGTA	1095
DB	198	AlaIleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyGluValSerPhe---	216
QY	1096	GCGAATGTCGAGAGTAGAGGGGGAGGATTTGCTGCTTTCAGGATGGGCAGCAGGA	1155
DB	216	-----	216
QY	1156	GTGTCATCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACTGGGTA	1215
DB	217	---SerAspAsnThrSerSerAspSer-----	224
QY	1216	GAGTTTGTGGAACGATAGCCCGAGTAGGAGGAGGATTTACTCTACGGGGAACGTTGCT	1275
DB	225	-----GlyAlaAlaIlePheThrGluAlaSerValThr	235
QY	1276	TTCTGTAATAATGGAATAACCTTCTTCTCAACAATGTTGCTTCTCTGTTTACATGCT	1335
DB	236	IleSerAsnAsnAlaLysValSerPheIleAspAsn-----	247
QY	1336	GCTAAGCAACCAACAGTGGACAGCTTCTTAATAGCAGTAATAATATTCGGAGATCGAGA	1395
DB	248	-----LysValThrGlyAlaSerSerThrThrGlyAspMet---SerGlyGly	263
QY	1396	GCTATCTTCTGTAAGAATGGTCGCAAGCAGGATCCAATAACTCTGGATCAGTTTCCCTTT	1455
DB	264	AlaIleCysAlaTyrLys-----ThrSerThrAspThrLysValThrLeu	278
QY	1456	GATGAGAGGAGTAGTATTTCTTTAGTAGCAATGATGCTGCGGAAAGGGGAGGACTATT	1515
DB	279	ThrGlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrAlaGlyAlaIle	298
QY	1516	TATGCCAAAAAGCTCTCGTTCCTAATCTGGCCCTGTACAATTTTAAAGGAATATCGCT	1575
DB	299	TyrValLysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsnSerVal	318
QY	1576	AAT-----GATGGTCGAGCGATTATTTAGGAGAAATCTGGAGAGCTCAGT	1620
DB	319	AsnGlyGlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGluLeuSer	338
QY	1621	TTATCTGCTGATTTATGGAGATATTATTTTCGATGGGAATCTTTAAAGAACACGCCAAGAG	1680
DB	339	LeuSerAlaAspSerGlyAspIleValPheLeuGlyAsn-----ThrValThrSer	355
QY	1681	AATGCTGGCGATTTAATGGCGTAATCTGTGCTCTCAAGCCATTTTCGATGGATCGGGA	1740
DB	356	ThrThrProGlyThrAsn-----ArgSerSerIleAspLeuGlyThrSer	370
QY	1741	GGGAAATAACCACTAATTAAGCTAAAGCAGGCGATCATGATTTCTTTTAAATGATCCATC	1800
DB	371	AlaLysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleThrPheTyrAspPheIle	390
QY	1801	GAGATGGCAACGGAAATAACAGCCAGCGCAGTCTTCCAAACTTCTTAAATAATTAACGAT	1860
DB	391	ThrThrGlySerSerThrThr-----ValThrAspValLeuLysValAsnGlu	406
QY	1861	GGTGAAGGA-----TACACAGGGGATATTGTTTTGCT-----	1893
DB	407	ThrProAlaAspSerAlaLeuGlnThrGlyAsnIleIlePheThrGlyGluLysLeu	426
QY	1894	-----AATGGAAGCAGTACTTTGTACCAAAATGTTTACG	1926
DB	427	SerGluThrGluAlaAlaAspSerLysAsnLeuThrSerLysLeuLeuGlnProValThr	446
QY	1927	ATAGAGCAAGGAGGATTTCTTCTGTAAGGCAAAATATATCAGTGAATTTCTTAAGT	1986
DB	447	LeuSerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThrGlnAlaPheThr	466
QY	1987	CAGACAGTGGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGAATCCCA	2043
DB	467	GlnGlnAlaAspSerArgLeuGluMetAspValGlyThrThrLeuGlu-----	482
QY	2044	CAACCAACCAACAGCTCTCGCGCTAATCAGTTGATCAGCTTTCATCTGCAATTCGATTG	2103
DB	483	-----ProAlaAspThrSer-----ThrIleAsnLeuValIle	494
QY	2104	TCCTTTCTTCTTTGTAGCAACAAATGCAATCTCTCAATCTCTCAATCTCCACGCG	2163
DB	495	AsnIleSerSerIle-----	499
QY	2164	CRAGATTCTCATCTCTCAGTCATTTGTAGC---ACAATGCTGCTGTTCTGTTACAATTAGT	2220
DB	500	AspGlyAlaLysLysAlaLysIleGluThrLysAlaThrSerLysAsnLeuThrLeuSer	519
QY	2221	GGCCTATCTTTTTCAGGATTTGGATGATACAGCTTATCATAGTAGTATGATTCGCTAGGT	2280
DB	520	GlyThrIleThrLeuLeuAspProThrGlyThrPheTyrGluAsnHisSer---LeuArg	538
QY	2281	TCTAATCAAAAAATCAATGTCCTGMAATTACAGTTA---GGGACTAAGACCCCGAGCTAAT	2337

```
Db 539 AsnProGlnSerTyrAspIleLeuGluLeuIysAlaSerGlyThrValThrSerThrAla 558
Qy 2338 GCCCATCAGATTGACCTCAGGGAATCAGATGCCAAGTATGGCTATCAAGGAGCTGG 2397
Db 559 ValThrProAspProIleMetGlyGluIysPhe---HisTyrGlyThrGlnGlyThrTrp 577
Qy 2398 ---AAGCTTGCGTGGATCCTAATACAGCAATAAATGGTCTCTTATATCTCTGAAGCTACA 2454
Db 578 GlyProIleValTrpGlyThrGlyAlaSerThrThrAlaThrPhe-----Asn 593
Qy 2455 TGGACTAAACCTGGGTATAAATCCTGGGCTGAGGAGTAGCTCTCTTGGTTCCTCAATAGT 2514
Db 594 TrpThrIysThrGlyTyrIleProAsnProGluArgIleGlySerLeuValProAsnSer 613
Qy 2515 TTATGGGATCATTTAGATATACGATCTGGCAATTCAGCAATTCAGCAAGTGTGTGAT 2574
Db 614 LeuTrpAsnAlaPheIleAspIleSerSerLeuHisTyrLeuMetGluThrAlaAsnGlu 633
Qy 2575 GGGCGCTCTTATGTCGAGGATATAGGTTTCTGGAGTTTCGAAATTTCTTATCATGAC 2634
Db 634 GlyLeuGlnGlyAspArgAlaPheTrpCysAlaGlyLeuSerAsnPheHisIysAsp 653
Qy 2635 CGCGATGCTTAGTCAGGATATCGGTATATAGTGGGGTTATCTCTAGGAGCAAC 2694
Db 654 SerThrIysThrArgArgGlyPheArgHisLeuSerGlyGlyTyrValIleGlyAsn 673
Qy 2695 TCCTACTTTGGATCA---TCGATGTTGCTAGCATTTACGAAATTTGGTAGATCT 2751
Db 674 LeuHisThrCysSerAspIysIleLeuSerAlaAlaPheCysGlnLeuPheGlyArgAsp 693
Qy 2752 AAAGATTATGATGTGCTGTTCCATCATCATGCTTCATAGGATCCGTTTATCTATCT 2811
Db 694 ArgAspTyrPheValAlaIysAsnGlnGlyThrValTyrGlyGlyThrLeuTyr-TyrG1 713
Qy 2812 ACCCAACAGCTTTATGTCGATCCCTATTGTCGGAGATGCGTTTATCGTGCT----- 2865
Db 713 nHisAsnGluThrTyrIleSerLeuProCysIysLeuArgProCysSerLeuSerTyrVa 733
Qy 2866 -----AGCTACGGGTTTGGAAATCAGCA 2888
Db 733 lProThrGluIleProValLeuPheSerGlyAsnLeuSerTyrThrHisThrAspAsnAs 753
Qy 2889 TATGAAAACCTCATATACATTTGCAGAGGAGCGATGTTCTGGGATAATACTGTCT 2948
Db 753 pLeuIysThrIysTyrThrTyrProThrValIysGlySerTrpGlyAsnAspSerPh 773
Qy 2949 GCCTGGAGATGTCAGCGGATACCGATTGTGATTACTCCTAAGCTCTATTGAA 3008
Db 773 eAlaLeuGluPheGlyGlyArgAlaProIleCysLeuAspGluSerAlaLeuPhe---G1 792
Qy 3009 TGAGTTGCGTCTTTCGTCAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGA 3068
Db 792 uGlnTyrMetProPheMetLeuIysLeuGlnPheValTyrAlaHisGlnGluGlyPheIysG1 812
Qy 3069 GGAAGCGCATCAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGT 3128
Db 812 uGlnGlyThrGluAlaArgGluPheGlySerSerArgLeuValAsnLeuAlaLeuProI1 832
Qy 3129 TGAGTGAAGTTTGATCGATGTTCTAGTACATCTTAATAATATAGCTTTATGCGCGC 3188
Db 832 eGlyIleArgPheAspLysGluSerAspCysGlnAspAlaThrTyrAsnLeuThrLeuG1 852
Qy 3189 TTATATCTGTGATGCTTATCGCACCATCTCTGTACTGAGACAACGCTCTCTATCCATCA 3248
Db 852 yTyrThrValAspLeuValArgSerAsnProAspCysThrThrThrLeuArgIleSerG1 872
Qy 3249 AGAGCATGACACACAGATGCTCTTTCATTTAGCAAGACATGAGGTGTGGTTAGGATC 3308
Db 872 yAspSerTyrIysThrPheGlyThrAsnLeuAlaArgGlnAlaLeuValLeuAlaG1 892
Qy 3309 TATGATGCTTCTTAACAGTAATATAGAAATATATGCCATGCGCATGAGATATGATTCG 3368
Db 3368 -----
```

```
Db 892 yAsnHisPheCysPheAsnSerAsnPheGluAlaPheSerGlnPheSerPheGluLeuAr 912
Qy 3369 AGATGCTTCTCGAGGCTATGTTTGTAGTCGACGAGTAGAGTCCGGTTC 3417
Db 912 gGlySerArgAsnTyrAsnValAspLeuGlyAlaLysTyrGlnPhe 928

RESULT 28
US-09-438-185A-453
; Sequence 453, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 453
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURES:
; OTHER INFORMATION: Cpn0451
US-09-438-185A-453

Alignment Scores:
Pred. No.: 2,39e-92 Length: 967
Score: 1094.50 Matches: 320
Percent Similarity: 42.6% Conservative: 157
Best Local Similarity: 28.6% Mismatches: 414
Query Match: 13.9% Indels: 229
DB: 2 Gaps: 32

US-10-701-844-1 (1-4435) x US-09-438-185A-453 (1-967)
Qy 286 AGGACGAATCGCTGCAAGATAAGA-----ACATTTATTGATATT 324
Db 1 LysThrGlnArgValIysIleLeuAspSerCysPheValIlePheAsnLeuIle 20
Qy 325 AAATATTAAATTTTATGAACGGAGTAATTAATTTATCTCTCAGCTTTGTGTGATG 384
Db 21 TyrLeuPheCysPheTyrIleAspAlaAsnSerSerLeuLysAsnLysSerIleThrMet 40
Qy 395 CAACGCTCTTCCATAAGTCTTCTTCAATGATCTTAGCTTATTCTTTC-----TGC 438
Db 41 LysThrSerIleProTrpValLeuValSerSerValLeuAlaPheSerCysHisLeuGln 60
Qy 439 TCTTTAAATGGGGGGGATATGCAGCAAAATCATGGTTCCTCAAGGAATTTACGATGG 498
Db 61 SerLeuAlaAsn-----GluGluLeuLeuSerProAspAspSerPheAsnGly 76
Qy 499 GAGCGTTAACTGATCATTT-----CCCTATATCTGTATAGGA 537
Db 77 AsnIleAspSerGlyThrPheThrProLysThrSerAlaThrThrTyrSerLeuThrGly 96
Qy 538 GATCCGAGTGGGACTACTGTTTTCCTGAGGAGAGTAACTTAATAAAAAATCTTGACAA 597
Db 97 Asp-----ValPhePheTyrGluProGlyLysGlyThr----- 107
Qy 598 TCTATTGCACTTTCCTTTA-----AGTTGTTTTGGGAATTTATAGGGATTTTACT 651
Db 108 -----ProLeuSerAspSerCysPheLysGlnThrThrAspAsnLeuThr 122
Qy 652 GTTTTAGGAGGAGGACTCGTTGACTTTCGAGACATACGAGCTTCTACAAATGGGCA 711
```

Db	123	pheLeuGlyAsnGlyHisSerLeuThrPheGlyPheIleAspAlaGlyThrHisAlaGly	142
Qy	712	GCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGTTTTAAAGAATATCC	771
Db	143	AlaAlaSerThrThrAlaAsnLeuThrPheSerGlyPheSerLeuLeuSer	162
Qy	772	TTTTCCAAATGCAATTCAATTACTTGCCTACTGCTGCTCAACGACTAATAAGGGTAGC	831
Db	163	PheAspSer	165
Qy	832	CAGACTCCGACGACAACATCTACACGCTCTAATGTAAGTACTATTTATCTAAAAACAGATCTT	891
Db	166	---SerProSerThrThrValThrThrGlyGlnGlyThrLeuSerSerAlaGlyGlyVal	184
Qy	892	TTGTACTCAATAATGAGAGTCTCATTTAGTAATTTAGTCTCTCGAGATGGGGA	951
Db	185	AsnLeuGluAsnIleArgLysLeuValAlaGlyAsnPheSerThrAlaAspGlyGly	204
Qy	952	GCTATAGATGCTAAGAGCTTAAACGGTTTCAAGGAATTAGCAAGCTTTGTGCTTCCAAGAA	1011
Db	205	AlaIleLysGlyAlaSerPheLeuLeuThrGlyThrSerGlyAspAlaLeuPheSerAsn	224
Qy	1012	AATACTGCTCAAGCTGATGGGGAGCTGTCAAGTAGTCAACAGTTTCTCTGCTGATGGCT	1071
Db	225	AsnSerSerSerThrLysGlyAla	233
Qy	1072	AACGAGGCTCTATTGCTTTGTATGCGAATTTGCAGAGCTAAGAGGGGAGGAGTTGCT	1131
Db	234	-----IleAlaThrAlaGlyAlaArg	241
Qy	1132	GCTGTTCCAGGATGGCAGCAGGAGTGTCTATCTACTTCAACAGAAGATCCAGTAGTA	1191
Db	241	-----	241
Qy	1192	AGTTTTCAGAAATATCTGCG---GTAGAGTTTGAATGGGAACGTAGCCGAGTAGGAGA	1248
Db	242	---IleAlaAsnAsnThrGlyTyThrValArgPheLeuSerAsnIleAlaSerThrSerGly	260
Qy	1249	GGGATTTACTCTCAGGGAAGTTCCTCTGAAATATGGAACAACTTGTCTTCTCAAC	1308
Db	261	GlyAlaIleAspAspGluGlyThrSerIleLeuSerAsnLysPheLeuTyPhe	279
Qy	1309	AATGTTGCTTCTCCTGTTTACATTGCTGCTAAGCAACCAAGTAGTGACAGGCTCTTAAT	1368
Db	280	-----GluGlyAsnAlaAlaLys	285
Qy	1369	ACGAGTAAATATACGAGATGAGGAGCTATCTTCTGTAAGAATGTTGGCAGCAGGA	1428
Db	286	ThrThr-----GlyGlyAlaIle---CysAsnThrLysAlaSerGlySer	299
Qy	1429	-----TCCAAATACTCGATCAGTTTCCTTGTATGAGAGGAGTAGTT	1473
Db	300	ProGluLeuIleIleSerAsnLys-----ThrLeu	310
Qy	1474	TTCTTTAGTACAAATGTAGCTGCTCGGAAGGGGAGCTATTTATGCCAAAAAGCTCTCG	1533
Db	311	IlePheAlaSerAsnValAlaGluThrSerGlyGlyAlaIleHisAlaLysLeuAla	330
Qy	1534	GTTGCTAACTGTGGCCCTGTACAATTTTAAAGAAATATCGCTAATGAT	1581
Db	331	LeuSerSerGlyGlyPheThrGluPheLeuArgAsnValSerSerAlaThrProLys	350
Qy	1582	GTTGAGCGATTTATTTAGGAGAACTCGAGAGCTCAGTTTATCTCTGATTTATGAGAT	1641
Db	351	GlyGlyAlaIleSerIleAspAlaSerGlyGluLeuSerLeuSerAlaGluThrGlyAsn	370
Qy	1642	ATTATTTTTCGATGGGAATCTTTAAAGAACACGCCAAGAGAAATGCTGCCGATGTTAATGCG	1701
Db	371	IleThrPheValArgAsnThrLeuThrThr-----GlySerThrAsp	385
Qy	1702	GTAAGTGTGCTTCAAGCCATTTTCGATGGGATCGGAGGGGAAAAATAACGACATTAAGA	1761

386	Db	---ThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLysPheThrGluLeuArg	404
1762	QY	GCTAAAGCAGGGCATCAGATTCTCTTTAAATGATPCCATCGAGATGCCAAACGGAAATAAC	1821
405	Db	AlaAlaLysAsnHisThrIlePhePheTyrAspProIle	418
1822	QY	CAGCAGCGCAGTCTTCCAAACTCTTAAATAATTAACAGATGGTCAAGGA	1869
419	Db	SerGluGlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeuAsn	438
1870	QY	---TACACAGGGGATATTGTTTTTGCT	1893
439	Db	ProTyrGlnGlyThrIleLeuPheSerGlyLeuThrLeuThrAlaAspGluLeuLysVal	458
1894	QY	---AATGGAAGCAGTACTTTGTATCCAAATGTTTACGATAGCAGCAAGAGGATGTT	1947
459	Db	AlaAspAsnLeuLysSerPheThrGlnProValSerLeuSerGlyGlyLysLeu	478
1948	QY	CTTCGTGAAGAAGCAAAATTTATCAGTGAATCTCTTAAGTCAGACAGGTGGAGCTCTGAT	2007
479	Db	LeuGlnLysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeu	498
2008	QY	---ATGGAAGCTGGAGTACATGGGATTTTGTAACTCCAACACACCACACAGCCCTCT	2064
499	Db	GlyMetAspSerGlyThrThrLeuSerThrThrAlaGlySer	512
2065	QY	GCCGCTAATCAGTTGATCAGCTTTCCATCTGCATTTGTCTCTTCTCTTTGTTAGCA	2124
513	Db	---IleThrIleThrAsnLeuGlyIleAsnValAspSerLeuGlyLeu	527
2125	QY	AACAATGCAGTTACGATCTCTCTACCAATCCTCCAGCGCAAGATTCTCATCTCGACATC	2184
528	Db	LysGlnProVal	533
2185	QY	ATTGCTAGCACAACTGCTGGTTCTGTTTCAATTAGTGGGCTATCTTTTGGAGGATTG	2244
534	Db	ThrAlaLysGlyAlaSerAsnLysValIleValSerGlyLysLeuAsnLeuIleAspIle	553
2245	QY	GATGATACAGCTTATGATAGTATGATGTGGCTAGTGTCTTAATCAAAAATCAATGTCCTG	2304
554	Db	GluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeuLeu	572
2305	QY	AAATTACAGTTAGGAGCTAAGCCCCAGCTAAATGCCCATCAGATTTTGACTCTAGGGAAT	2364
573	Db	LysIleThrVal	588
2365	QY	GAGATGCTCT	2400
589	Db	LeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlyGlnTrpAsn	608
2401	QY	CTTCGTGGGATCTTAATACAGCAAAATAATGGTCCTTATCTACTCGAAGCTACATGGACT	2460
609	Db	ValAsnTrpThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrpThr	627
2461	QY	AAAACTGGGTATATCTCGGCCCTGAGGAGTACTCTTTGGTTCCAAATAGTTTATGG	2520
628	Db	LysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeuTrp	647
2521	QY	GGATCCATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGCGC	2580
648	Db	GlyValPheThrAspIleArgSerLeuGlnGlnLeuValGluIleGlyAlaThrGlyMet	667
2581	QY	TCTTATTCGAGGATATAGGTTCTCGAGTTTCGAAATTTCTTCTATCATGACCGCAT	2640
668	Db	GluHisLysGlnGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGlyAsp	687
2641	QY	GCTTTAGTCAGGATATCGGTATATAGTCGGGGTTATTCTTTAGGAGCAATCTCTAC	2700
688	Db	GluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlyGlySerAlaHis	707
2701	QY	TTT---GGATCATCATGTTTGGTCTAGCATTTACCGAAGTATTTTGGTAGATCTAAAGAT	2757
708	Db	ThrProLysAspAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLysAsp	727



1246 GGAGGATTTACTCTACGGGAACGTTGCTTTCTGAATAATGGAACCAACCTGTTTCTC 1305  
Db ||||| : : : : : ||||| : : : : :  
242 GlyGlyAlaIleAspGluGlyThrSerIleLeuSerAsnAsnLysPheLeuTyrPhe 261  
Qy ||||| : : : : : ||||| : : : : :  
1306 AACAAATGTTCTCTCTCTGTTTACATTTGCTGCTAAACCAACCAAGTGGAGAGCTTCT 1365  
Qy ||||| : : : : : ||||| : : : : :  
262 -----GluGlyAsnAlaIle 266  
Db ||||| : : : : :  
1366 AATACGAGTAATAATACGGAGATCGAGGAGCTATCTTCTGTGAAGATGGTGGCGAACA 1425  
Qy ||||| : : : : : ||||| : : : : :  
267 LysThrThr-----GlyGlyAlaIle---CysAsnThrLysAlaSerGly 280  
Db ||||| : : : : : ||||| : : : : :  
1426 GGA-----TCCAATAACTCTGGATCAGTTTCTCTTGTGATGGAGAGGAGTA 1470  
Qy ||||| : : : : : ||||| : : : : :  
281 SerProGluLeuIleSerAsnLys-----Thr 291  
Db ||||| : : : : : ||||| : : : : :  
1471 GTTTCTTTAGTACCAATGAGTCTCTGGGAAAGGGAGCTATTATTATGCCAAAAGCTC 1530  
Qy ||||| : : : : : ||||| : : : : :  
292 LeuIlePheAlaSerAsnValAlaGluThrSerGlyAlaIleHisAlaLysLysLeu 311  
Db ||||| : : : : : ||||| : : : : :  
1531 TCGGTTGCTAACTGTGGCCCTGTACATTTTAAAGAAATATCGTAATGAT----- 1581  
Qy ||||| : : : : : ||||| : : : : :  
312 AlaLeuSerSerGlyGlyPheThrGluPheLeuArgAsnAsnValSerSerAlaThrPro 331  
Db ||||| : : : : : ||||| : : : : :  
1582 ---GGTGGAGCATTTATTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTATGA 1638  
Qy ||||| : : : : : ||||| : : : : :  
332 LysGlyGlyAlaIleSerIleAspAlaSerGlyGlyLeuSerLeuSerAlaGluThrGly 351  
Db ||||| : : : : : ||||| : : : : :  
1639 GATATTATTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCCGATGTTAAT 1698  
Qy ||||| : : : : : ||||| : : : : :  
352 AsnIleThrPheValArgAsnThrLeuThrThr-----GlySerThrAsp----- 367  
Db ||||| : : : : : ||||| : : : : :  
1699 GCGGTAACTGTCTCTCAACAGCATTTCTGATGGGATCGGGAGGGAGAAATAAGCATTA 1758  
Qy ||||| : : : : : ||||| : : : : :  
368 -----ThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLysPheThrGluLeu 385  
Db ||||| : : : : : ||||| : : : : :  
1759 AGAGCTAAAGCAGGCGATCTCTTTTAAATGATCCCATCGAGATGCGCAACGGAAT 1818  
Qy ||||| : : : : : ||||| : : : : :  
386 ArgAlaAlaLysAsnHisThrIlePhePheTyrAspProIle----- 399  
Db ||||| : : : : : ||||| : : : : :  
1819 AACACGCGCAGCTCTTCCAACTTCTAAAATTAACGATGCTGAAGGA----- 1869  
Qy ||||| : : : : : ||||| : : : : :  
400 ThrSerGluGlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeu 419  
Db ||||| : : : : : ||||| : : : : :  
1870 -----TACACAGGGATATGTTTGTCT----- 1893  
Qy ||||| : : : : : ||||| : : : : :  
420 AsnProTyrGlnGlyThrIleLeuPheSerGlyGluThrLeuThrAlaAspGluLeuLys 439  
Db ||||| : : : : : ||||| : : : : :  
1894 -----AATGGAAGCAGTACTTGTACCAAAATGTTACGATAGCAGCAAGAGGATT 1944  
Qy ||||| : : : : : ||||| : : : : :  
440 ValAlaAspAsnLeuLysSerPheThrGlnProValSerLeuSerGlyGlyLysLeu 459  
Db ||||| : : : : : ||||| : : : : :  
1945 GTTCTTCTGTAAGGCAAAATATCAGTGAATCTCTAAGTCAGACAGGTGGAGCTCG 2004  
Qy ||||| : : : : : ||||| : : : : :  
460 LeuLeuGlnIysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeu 479  
Db ||||| : : : : : ||||| : : : : :  
2005 TAT---ATGGAAGCTCGGAGTACATGGGATTTTGTAACTCCACCAACCAACAGCCT 2061  
Qy ||||| : : : : : ||||| : : : : :  
480 LeuGlyMetAspSerGlyThrThrLeuSerThrThrAlaGlySer----- 494  
Db ||||| : : : : : ||||| : : : : :  
2062 CTTGCCGCTAATCAGTTGATCAGCTTCTCAATCTGCATTTGTCTCTTCTTTTGTGA 2121  
Qy ||||| : : : : : ||||| : : : : :  
495 -----IleThrIleThrAsnLeuGlyIleAsnValAspSerLeuGly 508  
Db ||||| : : : : : ||||| : : : : :  
2122 GCAACAAATGCAGTTACGAATCTCTACCAATCTCCAGCGCAAGATTCTCATCTGCA 2181  
Qy ||||| : : : : : ||||| : : : : :  
509 LeuLysGlnProVal-----Ser 514  
Db ||||| : : : : : ||||| : : : : :  
2182 GTCATTGGTAGCAACAACCTGCTGTTCTGTATACAAATTAGTGGGCTATCTTTTGGAGAT 2241  
Qy ||||| : : : : : ||||| : : : : :  
515 LeuThrAlaLysGlyAlaSerAsnLysValIleValSerGlyLysLeuAsnLeuIleAsp 534  
Db ||||| : : : : : ||||| : : : : :  
2242 TTGGATGATACAGCTTATGATAGTATGATGGCTAGTGTCTAATCAAAAATCAATGTC 2301  
Qy ||||| : : : : : ||||| : : : : :  
535 IleGluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553  
Qy ||||| : : : : : ||||| : : : : :  
2302 CTGAAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTAGGG 2361  
Db ||||| : : : : : ||||| : : : : :  
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569  
Qy ||||| : : : : : ||||| : : : : :  
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397  
Qy ||||| : : : : : ||||| : : : : :  
570 SerLeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlnTyr 589  
Db ||||| : : : : : ||||| : : : : :  
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457  
Qy ||||| : : : : : ||||| : : : : :  
590 AsnValAsnTyrThrThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTyr 608  
Db ||||| : : : : : ||||| : : : : :  
2458 ACTAAAACCTGGTATAATCTCTGGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517  
Qy ||||| : : : : : ||||| : : : : :  
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628  
Db ||||| : : : : : ||||| : : : : :  
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577  
Qy ||||| : : : : : ||||| : : : : :  
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648  
Db ||||| : : : : : ||||| : : : : :  
2578 CGCTCTTATTGTTCGAGGATATGCTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637  
Qy ||||| : : : : : ||||| : : : : :  
649 MetGluHisLysGlnGlyPheTyrValSerSerMetThrAsnPheLeuHisLysThrGly 668  
Db ||||| : : : : : ||||| : : : : :  
2638 GATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATCTCTAGGACCAACTCC 2697  
Qy ||||| : : : : : ||||| : : : : :  
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlySerAla 688  
Db ||||| : : : : : ||||| : : : : :  
2698 TACTTT---GGATCATCGATGTTTCTGTAGCATTTACCGAAGTATTTGGTAGATCTAAA 2754  
Qy ||||| : : : : : ||||| : : : : :  
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708  
Db ||||| : : : : : ||||| : : : : :  
2755 GATTATGATGATGTCGTTTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814  
Qy ||||| : : : : : ||||| : : : : :  
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrGlyGlyThrLeuPhePheLysHis 728  
Db ||||| : : : : : ||||| : : : : :  
2815 CAACAAGCTTTA-----TGTGGATCTCTTTTGTTCGGAGATCGC 2853  
Qy ||||| : : : : : ||||| : : : : :  
729 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyAlaLysPheSerGluSer 748  
Db ||||| : : : : : ||||| : : : : :  
2854 TTTATC-----CGTGTAGCTACGGG 2874  
Qy ||||| : : : : : ||||| : : : : :  
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768  
Db ||||| : : : : : ||||| : : : : :  
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCGCATGTTCTGGTGG 2934  
Qy ||||| : : : : : ||||| : : : : :  
769 HisSerAspAsnArgMetGluThrHisTyrThrSerLeuProGluSerGluGlySerTyr 788  
Db ||||| : : : : : ||||| : : : : :  
2935 GATAATACTGTCTGCTGGAGATGAGCGGGATTTACCGATTTGATGTTGATTACTCTCT 2994  
Qy ||||| : : : : : ||||| : : : : :  
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808  
Db ||||| : : : : : ||||| : : : : :  
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGGAGCTGAGTTTCTTATCGGATCAT 3054  
Qy ||||| : : : : : ||||| : : : : :  
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrValSerGln 828  
Db ||||| : : : : : ||||| : : : : :  
3055 GAATCTTTTACAGAGGAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTAAAT 3114  
Qy ||||| : : : : : ||||| : : : : :  
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848  
Db ||||| : : : : : ||||| : : : : :  
3115 CTATCAGTTCTGTTGGAGTGAAGTTTGTATCGATGTTGTAGTACATCATCTTAATAATAT 3174  
Qy ||||| : : : : : ||||| : : : : :  
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyrThrTyr 868  
Db ||||| : : : : : ||||| : : : : :  
3175 AGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTGAGACAACG 3234  
Qy ||||| : : : : : ||||| : : : : :  
869 AspLeuSerGlyPhePheValSerAspValTyrArgAsnAsnProGlnSerThrAlaThr 888  
Db ||||| : : : : : ||||| : : : : :  
3235 CTCTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATTTAGCAACATGGATT 3294  
Qy ||||| : : : : : ||||| : : : : :  
535 IleGluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553  
Qy ||||| : : : : : ||||| : : : : :  
2302 CTGAAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTAGGG 2361  
Db ||||| : : : : : ||||| : : : : :  
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569  
Qy ||||| : : : : : ||||| : : : : :  
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397  
Qy ||||| : : : : : ||||| : : : : :  
570 SerLeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlnTyr 589  
Db ||||| : : : : : ||||| : : : : :  
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457  
Qy ||||| : : : : : ||||| : : : : :  
590 AsnValAsnTyrThrThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTyr 608  
Db ||||| : : : : : ||||| : : : : :  
2458 ACTAAAACCTGGTATAATCTCTGGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517  
Qy ||||| : : : : : ||||| : : : : :  
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628  
Db ||||| : : : : : ||||| : : : : :  
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577  
Qy ||||| : : : : : ||||| : : : : :  
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648  
Db ||||| : : : : : ||||| : : : : :  
2578 CGCTCTTATTGTTCGAGGATATGCTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637  
Qy ||||| : : : : : ||||| : : : : :  
649 MetGluHisLysGlnGlyPheTyrValSerSerMetThrAsnPheLeuHisLysThrGly 668  
Db ||||| : : : : : ||||| : : : : :  
2638 GATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATCTCTAGGACCAACTCC 2697  
Qy ||||| : : : : : ||||| : : : : :  
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlySerAla 688  
Db ||||| : : : : : ||||| : : : : :  
2698 TACTTT---GGATCATCGATGTTTCTGTAGCATTTACCGAAGTATTTGGTAGATCTAAA 2754  
Qy ||||| : : : : : ||||| : : : : :  
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708  
Db ||||| : : : : : ||||| : : : : :  
2755 GATTATGATGATGTCGTTTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814  
Qy ||||| : : : : : ||||| : : : : :  
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrGlyGlyThrLeuPhePheLysHis 728  
Db ||||| : : : : : ||||| : : : : :  
2815 CAACAAGCTTTA-----TGTGGATCTCTTTTGTTCGGAGATCGC 2853  
Qy ||||| : : : : : ||||| : : : : :  
729 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyAlaLysPheSerGluSer 748  
Db ||||| : : : : : ||||| : : : : :  
2854 TTTATC-----CGTGTAGCTACGGG 2874  
Qy ||||| : : : : : ||||| : : : : :  
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768  
Db ||||| : : : : : ||||| : : : : :  
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCGCATGTTCTGGTGG 2934  
Qy ||||| : : : : : ||||| : : : : :  
769 HisSerAspAsnArgMetGluThrHisTyrThrSerLeuProGluSerGluGlySerTyr 788  
Db ||||| : : : : : ||||| : : : : :  
2935 GATAATACTGTCTGCTGGAGATGAGCGGGATTTACCGATTTGATGTTGATTACTCTCT 2994  
Qy ||||| : : : : : ||||| : : : : :  
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808  
Db ||||| : : : : : ||||| : : : : :  
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGGAGCTGAGTTTCTTATCGGATCAT 3054  
Qy ||||| : : : : : ||||| : : : : :  
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrValSerGln 828  
Db ||||| : : : : : ||||| : : : : :  
3055 GAATCTTTTACAGAGGAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTAAAT 3114  
Qy ||||| : : : : : ||||| : : : : :  
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848  
Db ||||| : : : : : ||||| : : : : :  
3115 CTATCAGTTCTGTTGGAGTGAAGTTTGTATCGATGTTGTAGTACATCATCTTAATAATAT 3174  
Qy ||||| : : : : : ||||| : : : : :  
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyrThrTyr 868  
Db ||||| : : : : : ||||| : : : : :  
3175 AGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTGAGACAACG 3234  
Qy ||||| : : : : : ||||| : : : : :  
869 AspLeuSerGlyPhePheValSerAspValTyrArgAsnAsnProGlnSerThrAlaThr 888  
Db ||||| : : : : : ||||| : : : : :  
3235 CTCTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATTTAGCAACATGGATT 3294  
Qy ||||| : : : : : ||||| : : : : :  
535 IleGluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553  
Qy ||||| : : : : : ||||| : : : : :  
2302 CTGAAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTAGGG 2361  
Db ||||| : : : : : ||||| : : : : :  
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569  
Qy ||||| : : : : : ||||| : : : : :  
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397  
Qy ||||| : : : : : ||||| : : : : :  
570 SerLeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlnTyr 589  
Db ||||| : : : : : ||||| : : : : :  
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457  
Qy ||||| : : : : : ||||| : : : : :  
590 AsnValAsnTyrThrThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTyr 608  
Db ||||| : : : : : ||||| : : : : :  
2458 ACTAAAACCTGGTATAATCTCTGGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517  
Qy ||||| : : : : : ||||| : : : : :  
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628  
Db ||||| : : : : : ||||| : : : : :  
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577  
Qy ||||| : : : : : ||||| : : : : :  
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648  
Db ||||| : : : : : ||||| : : : : :  
2578 CGCTCTTATTGTTCGAGGATATGCTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637  
Qy ||||| : : : : : ||||| : : : : :  
649 MetGluHisLysGlnGlyPheTyrValSerSerMetThrAsnPheLeuHisLysThrGly 668  
Db ||||| : : : : : ||||| : : : : :  
2638 GATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATCTCTAGGACCAACTCC 2697  
Qy ||||| : : : : : ||||| : : : : :  
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlySerAla 688  
Db ||||| : : : : : ||||| : : : : :  
2698 TACTTT---GGATCATCGATGTTTCTGTAGCATTTACCGAAGTATTTGGTAGATCTAAA 2754  
Qy ||||| : : : : : ||||| : : : : :  
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708  
Db ||||| : : : : : ||||| : : : : :  
2755 GATTATGATGATGTCGTTTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814  
Qy ||||| : : : : : ||||| : : : : :  
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrGlyGlyThrLeuPhePheLysHis 728  
Db ||||| : : : : : ||||| : : : : :  
2815 CAACAAGCTTTA-----TGTGGATCTCTTTTGTTCGGAGATCGC 2853  
Qy ||||| : : : : : ||||| : : : : :  
729 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyAlaLysPheSerGluSer 748  
Db ||||| : : : : : ||||| : : : : :  
2854 TTTATC-----CGTGTAGCTACGGG 2874  
Qy ||||| : : : : : ||||| : : : : :  
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768  
Db ||||| : : : : : ||||| : : : : :  
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCGCATGTTCTGGTGG 2934  
Qy ||||| : : : : : ||||| : : : : :  
769 HisSerAspAsnArgMetGluThrHisTyrThrSerLeuProGluSerGluGlySerTyr 788  
Db ||||| : : : : : ||||| : : : : :  
2935 GATAATACTGTCTGCTGGAGATGAGCGGGATTTACCGATTTGATGTTGATTACTCTCT 2994  
Qy ||||| : : : : : ||||| : : : : :  
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808  
Db ||||| : : : : : ||||| : : : : :  
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGGAGCTGAGTTTCTTATCGGATCAT 3054  
Qy ||||| : : : : : ||||| : : : : :  
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrValSerGln 828  
Db ||||| : : : : : ||||| : : : : :  
3055 GAATCTTTTACAGAGGAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTAAAT 3114  
Qy ||||| : : : : : ||||| : : : : :  
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848  
Db ||||| : : : : : ||||| : : : : :  
3115 CTATCAGTTCTGTTGGAGTGAAGTTTGTATCGATGTTGTAGTACATCATCTTAATAATAT 3174  
Qy ||||| : : : : : ||||| : : : : :  
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyrThrTyr 868  
Db ||||| : : : : : ||||| : : : : :  
3175 AGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTGAGACAACG 3234  
Qy ||||| : : : : : ||||| : : : : :  
869 AspLeuSerGlyPhePheValSerAspValTyrArgAsnAsnProGlnSerThrAlaThr 888  
Db ||||| : : : : : ||||| : : : : :  
3235 CTCTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATTTAGCAACATGGATT 3294  
Qy ||||| : : : : : ||||| : : : : :  
535 IleGluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553  
Qy ||||| : : : : : ||||| : : : : :  
2302 CTGAAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTAGGG 2361  
Db ||||| : : : : : ||||| : : : : :  
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569  
Qy ||||| : : : : : ||||| : : : : :  
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397  
Qy ||||| : : : : : ||||| : : : : :  
570 SerLeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlnTyr 589  
Db ||||| : : : : : ||||| : : : : :  
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457  
Qy ||||| : : : : : ||||| : : : : :  
590 AsnValAsnTyrThrThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTyr 608  
Db ||||| : : : : : ||||| : : : : :  
2458 ACTAAAACCTGGTATAATCTCTGGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517  
Qy ||||| : : : : : ||||| : : : : :  
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628  
Db ||||| : : : : : ||||| : : : : :  
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577  
Qy ||||| : : : : : ||||| : : : : :  
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648  
Db ||||| : : : : : ||||| : : : : :  
2578 CGCTCTTATTGTTCGAGGATATGCTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637  
Qy ||||| : : : : : ||||| : : : : :  
649 MetGluHisLysGlnGlyPheTyrValSerSerMetThrAsnPheLeuHisLysThrGly 668  
Db ||||| : : : : : ||||| : : : : :  
2638 GATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATCTCTAGGACCAACTCC 2697  
Qy ||||| : : : : : ||||| : : : : :  
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlySerAla 688  
Db ||||| : : : : : ||||| : : : : :  
2698 TACTTT---GGATCATCGATGTTTCTGTAGCATTTACCGAAGTATTTGGTAGATCTAAA 2754  
Qy ||||| : : : : : ||||| : : : : :  
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708  
Db ||||| : : : : : ||||| : : : : :  
2755 GATTATGATGATGTCGTTTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814  
Qy ||||| : : : : : ||||| : : : : :  
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrGlyGlyThrLeuPhePheLysHis 728  
Db ||||| : : : : : ||||| : : : : :  
2815 CAACAAGCTTTA-----TGTGGATCTCTTTTGTTCGGAGATCGC 2853  
Qy ||||| : : : : : ||||| : : : : :  
729 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyAlaLysPheSerGluSer 748  
Db ||||| : : : : : ||||| : : : : :  
2854 TTTATC-----CGTGTAGCTACGGG 2874  
Qy ||||| : : : : : ||||| : : : : :  
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768  
Db ||||| : : : : : ||||| : : : : :  
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCGCATGTTCTGGTGG 2934  
Qy ||||| : : : : : ||||| : : : : :  
769 HisSerAspAsnArgMetGluThrHisTyrThrSerLeuProGluSerGluGlySerTyr 788  
Db ||||| : : : : : ||||| : : : : :  
2935 GATAATACTGTCTGCTGGAGATGAGCGGGATTTACCGATTTGATGTTGATTACTCTCT 2994  
Qy ||||| : : : : : ||||| : : : : :  
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808  
Db ||||| : : : : : ||||| : : : : :  
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGGAGCTGAGTTTCTTATCGGATCAT 3054  
Qy ||||| : : : : : ||||| : : : : :  
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrValSerGln 828  
Db ||||| : : : : : ||||| : : : : :  
3055 GAATCTTTTACAGAGGAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTAAAT 3114  
Qy ||||| : : : : : ||||| : : : : :  
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848  
Db ||||| : : : : : ||||| : : : : :  
3115 CTATCAGTTCTGTTGGAGTGAAGTTTGTATCGATGTTGTAGTACATCATCTTAATAATAT 3174  
Qy ||||| : : : : : ||||| : : : : :  
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyrThrTyr 868  
Db ||||| : : : : : ||||| : : : : :  
3175 AGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTGAGACAACG 3234  
Qy ||||| : : : : : ||||| : : : : :  
869 AspLeuSerGlyPhePheValSerAspValTyrArgAsnAsnProGlnSerThrAlaThr 888  
Db ||||| : : : : : ||||| : : : : :  
3235 CTCTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATTTAGCAACATGGATT 3294  
Qy ||||| : : : : : ||||| : : : : :  
535 IleGluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553  
Qy ||||| : : : : : ||||| : : : : :  
2302 CTGAAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTAGGG 2361  
Db ||||| : : : : : ||||| : : : : :  
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569  
Qy ||||| : : : : : ||||| : : : : :  
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397  
Qy ||||| : : : : : ||||| : : : : :  
570 SerLeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlnTyr 589  
Db ||||| : : : : : ||||| : : : : :  
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457  
Qy ||||| : : : : : ||||| : : : : :  
590 AsnValAsnTyrThrThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTyr 608  
Db ||||| : : : : : ||||| : : : : :  
2458 ACTAAAACCTGGTATAATCTCTGGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517  
Qy ||||| : : : : : ||||| : : : : :  
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628  
Db ||||| : : : : : ||||| : : : : :  
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577  
Qy ||||| : : : : : ||||| : : : : :  
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648  
Db ||||| : : : : : ||||| : : : : :  
2578 CGCTCTTATTGTTCGAGGATATGCTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637  
Qy ||||| : : : : : ||||| : : : : :  
649 MetGluHisLysGlnGlyPheTyrValSerSerMetThrAsnPheLeuHisLysThrGly 668  
Db ||||| : : : : : ||||| : : : : :  
2638 GATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATCTCTAGGACCAACTCC 2697  
Qy ||||| : : : : : ||||| : : : : :  
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlySerAla 688  
Db ||||| : : : : : ||||| : : : : :  
2698 TACTTT---GGATCATCGATGTTTCTGTAGCATTTACCGAAGTATTTGGTAGATCTAAA 2754  
Qy ||||| : : : : : ||||| : : : : :  
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708  
Db ||||| : : : : : ||||| : : : : :  
2755 GATTATGATGATGTCGTTTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814  
Qy ||||| : : : : : ||||| : : : : :  
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrGlyGlyThrLeuPhePheLysHis 728  
Db ||||| : : : : : ||||| : : : : :  
2815 CAACAAGCTTTA-----TGTGGATCTCTTTTGTTCGGAGATCGC 2853  
Qy ||||| : : : : : ||||| : : : : :  
729 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyAlaLysPheSerGluSer 748  
Db ||||| : : : : : ||||| : : : : :  
2854 TTTATC-----CGTGTAGCTACGGG 2874  
Qy ||||| : : : : : ||||| : : : : :  
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspVal

```
Db      889  LeuValMetSerProAepSerTrpLysIleArgGlyAenLeuSerArgGlnAlaPhe 908
Qy      3295 GTGCTTATGAGGATCTATGATCTCTCTAACAAGTAATATAGAAGTATATGGCCATGGA 3354
Db      909  LeuLeuArgGlySerAenAenTyrValTyrAenSerAenCysGluLeuPheGlyHisTyr 928
Qy      3355 AGATATAGATATCAGATGCTCTCGAGGCTATGTTGATGTCGAGGAGTAGAGTCCGG 3414
Db      929  AlaMetGluLeuArgGlySerSerArgAenTyrAenValAepValGlyThrLysLeuArg 948
Qy      3415 TTC 3417
Db      949  Phe 949

RESULT 30
US-09-198-452A-470
; Sequence 470, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 470
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-470

Alignment Scores:
Pred. No.:      2,73e-91      Length:      930
Score:          1083.00      Matches:     304
Percent Similarity: 43.3%      Conservative: 160
Best Local Similarity: 28.4%      Mismatches:  406
Query Match:     13.7%      Indels:      202
DB:              2          Gaps:         29

US-10-701-844-1 (1-4435) x US-09-198-452A-470 (1-930)
Qy      382  ATGCAAGCTCTTCCATAAGTCTCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT 441
Db      1  MetLysIleProLeuHisLysLeuLeuIleSerSerThrLeuVal---ThrProIleLeu 19
Qy      442  TTAATGGGGGGATATGCAGCAAAATC---ATGGTTCTCCTCAGGAAATTCAGATGG 498
Db      20  LeuSerIleAlaThrTyrGlyAlaAepAlaSerLeuSerProThrAepSerPheAepGly 39
Qy      499  GAGAGCTTAACGTATCATCTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTT 558
Db      40  AlaGlyLysSerThrPheThrProLysSerThr---AlaAepAlaAenGlyThrAenTyr 58
Qy      559  TTTCTCTGCGAGGAGTAAACATTAATAAAATCTTGACAAATCTATTGCGAGCTTTGCGCTTA 618
Db      59  ValLeuSerGlyAenValTyrIleAenAepAlaGlyLysGly---ThrAlaLeuThrGly 77
Qy      619  AGTTGTTTGGAACTTATAGGAGTCTTCTGTTTATAGGAGGAGGACACTCGTTGACT 678
Db      78  CysCysPheThrGluThrThrGlyAepLeuThrPheThrGlyLysGlyTyrSerPheSer 97
Qy      679  TTCGAGACATACCGACTTCTCAAAATGGGCGAGCTCAAGTAATAGGCTGCTCATGGA 738
Db      98  PheAenThrValAepAlaGlySerAenAlaGlyAlaAalaserThrThrAlaAepLys 117
Qy      739  CTGTTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTCTGCC 798
Db      118  AlaLeuThrPheThrGlyPheSerAenLeuSerPhe-----11e 130
Qy      799  GTACTGCTGCTGCAACGACTAATAAGGTTAGCGACTCCGACGACACA-----849
```

```
Db      131  AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu 150
Qy      850  TCTACACCGCTAATGCTACTATTTATTTCTTAAACAGAGATCTTTTGTACTCAATAATGAG 909
Db      151  AenLeuThrAepAenGlyThrIleLeuPheSerGlnAenVal-----164
Qy      910  AAGTTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGC 969
Db      165  -----SerAenGluAlaAenAenAenGlyGlyAlaIleThrAlaLysThr 179
Qy      970  TTAACGGTTCAGGAATTAGCAAGCTTTGTGCTCTTCAAGAAATATCTGCTCAAGCTGAT 1029
Db      180  LeuSerIleSerGlyAenThrSerSerIleThrPheThrSerAenSerAlaLysLysLeu 199
Qy      1030  GGGGAGCTTGTCAAGTAGTCACCGAGTTCTCTGCTATGCTAAGCTAAGAGGCTCTTATGCC 1089
Db      200  GlyGlyAlaIleTyrSerSerAlaAlaAalaserIleSerGlyAenThrGlyGlnLeuVal 219
Qy      1090  TTTGTAGGAATGTTGCGAGGTAAGAGGGAGGAGGATTCGTCTGTTTCAGGATGGGCAG 1149
Db      220  PheMetAenAen---LysGlyGluThrGlyGlyAlaLeuGlyPheGlu-----235
Qy      1150  CAGGAGGTGTCATCATCTCTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACT 1209
Db      236  -----AlaSerSerSerIleThrGlnAenSerSerLeuPhePheSerGlyAenThr 252
Qy      1210  GCGGTAGAGTTTGATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCAGGGAAC 1269
Db      253  AlaThrAep-----255
Qy      1270  GTTCTCTTCCTGAATTAATGGAAAAACCTTGTCTCAACATGTTGCTTCTCTGTTTAC 1329
Db      255  -----255
Qy      1330  ATTGCTGCTAAGCAACCAAGTGGACAGGCTTCTTAATACGAGTAATAATTACGAGAT 1389
Db      256  ---AlaAla-----GlyLys 259
Qy      1390  GGAGGAGCTATCTTCTGTAAGAATGGTGGCAAGCAGGATCCCAATAACTCTGGATCAGTT 1449
Db      260  GlyGlyAlaIleTyrCysGluLys-----ThrGlyGluThrProThrLeu 274
Qy      1450  TCCTTTGATGAGAGGAGTAGTCTTCTTCTAGTCAATGATGCTGCTGGGAAGGGGA 1509
Db      275  ThrIleSerGlyAenLysSerLeuThrPheAlaGluAenSerSerValThrGlnGly 294
Qy      1510  GCTATTTATGCCAAAAAGCTCTCGTTCCTAATCTGTCCTGTACAAATTTTAAAGAAAT 1569
Db      295  AlaIleCysAlaHisGlyLysAepSerAlaAlaGlyProThrLeuPheSerAenAen 314
Qy      1570  -----ATCGCTAATGATGGTGGAGCGATTTATTTAGGAGAATCTCGAGAG 1614
Db      315  ArgCysGlyAenThrAlaAlaGlyLysGlyGlyAlaIleAlaIleAlaAepSerGlySer 334
Qy      1615  CTCAGTTTATCTGCTGATATGAGATATTTATTTTCGATGGGAATCTTAAAGAAGACGCC 1674
Db      335  LeuSerLeuSerAlaAenGlnGlyAepIleThrPheLeuGlyAen-----349
Qy      1675  AAAGAGATGCTGCCGATTTAATGGCGTAACTGTCTCACAAGCCATTTTCGATGGA 1734
Db      350  -----ThrLeuThrSerThrSerAlaProThrSerThrArgAenAlaIleTyrLeuGly 367
Qy      1735  TCGGAGGAGGAAAAATACGACATTTAAGAGCTAAAGCAGGCGCATCAGATCTCTTTAATGAT 1794
Db      368  SerSerAlaLysIleThrAenLeuArgAlaAlaGlnGlyGlnSerIleTyrPheTyrAep 387
Qy      1795  CCCATCGAGATGGCAACCGAAAT-----AACGAGCCAGCGCAG 1833
Db      388  ProIleAlaSerAenThrThrGlyAlaSerAepValLeuThrIleAenGlnProAepSer 407
Qy      1834  TCTTCCAAACTTCTTAAAAATTAACGATGTTGAAGGATACACAGGCGGATTTCTTTTGTCT 1893
Db      408  AenSerProLeu-----AspTyrSerGlyThrIleValPheSer 420
```

```

1894  QY 1984 -----AATGGAAGCAGTACTTCTGAC 1914
Db 421 GlyGluLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys 440
QY 1915 CAAATATGTTACGATAGAGCAAGAGGATTGTTCTCGTGAAGAAAGCAAAATATACAGTG 1974
Db 441 GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLeuLysGlyAsnValGluLeuAspVal 460
QY 1975 AATTCTCTAAGTCAGACAGTGGAGTCTGTATATATGAAGCTGGGAGTACATGGGATTTT 2034
Db 461 AsnGlyPheThrGlnThrGluGlySerThrLeuLeu----- 472
QY 2035 GTAACCTCCAAACCAACCAACAGACCTCTCGCGCTAATCATGTTGATCAGCTTTCCAAAT 2094
Db 473 -----MetGlnProGlyThrLysLeuLysAlaAspThrGluAlaIleSerLeuThrLys 490
QY 2095 CTGCATTGCTCTTCTTCTTCTTGTAGCAACATGTCAGTGTACGAATCCTCTACCAAT 2154
Db 491 LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer----- 505
QY 2155 CCTCCAGCAGCAAGATTCTCATCTCGCAGTCATTGGGTAGCACAACTGCTGGT----- 2205
Db 506 -----IleGluThrAlaGlyAlaAsnLys 513
QY 2206 TCTGTTACAAATTAGTGGGCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGG 2265
Db 514 ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheThrGlu----- 532
QY 2266 TATGATTGGCTAGTTCTTAATCAAAAATCAATGCTCTGAAATTACAGTTAGGAGCTAAG 2325
Db 533 -----SerHisThrIleAsn-----GlnAlaPheThrGln 542
QY 2326 CCC-----CCAGCTAATGCCCCATCAGATTTCATCTAGGG----- 2361
Db 543 ProLeuValValPheThrAlaAlaThrAlaAlaSerAspIleThrIleAspAlaLeuLeu 562
QY 2362 -----AATGAGATGCCTAAGTATGGCTATCAGCAAGAGCTGGAAGCTGGG 2406
Db 563 ThrSerProValGlnThrProGluProHisThrGlyThrGlnGlyHisThrProGluAlaThr 582
QY 2407 TGG---GATCCTAATACAGCAATATGTCCTTATCTGAAAGCTACATCGACTAAA 2463
Db 583 TrpAlaAspThrSerThrAlaLysSerGly-----ThrMetThrTrpValThr 598
QY 2464 ACTGGGTATATCTCTGGCCTGAGCGAGTACTTCTTGTGTTCCAAATAGTTTATGGGA 2523
Db 599 ThrGlyThrAsnProAsnProGluAlaGArgAlaSerValProAspSerLeuTrpAla 618
QY 2524 TCCATTTTATAGATATACATCTCGGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCT 2583
Db 619 SerPheThrAspIleArgThrLeuGlnGlnIleMetThrSerGlnAlaAsnSerIleThr 638
QY 2584 TATGTCGAGGATATGGGTTCTCGGAGTTTCGAATTTCTTCTATCATGACCGCGATGCT 2643
Db 639 GlnGlnArgGlyLeuTrpAlaSerGlyThrAlaAsnPheHisLysAspLysSerGly 658
QY 2644 TTAGGTCAGGATATCGGTATATTAGTGGGGTTTATCTCTTAGGAGCAAACTCCTAC--- 2700
Db 659 ThrAsnGlnAlaPheArgHisLysSerThrGlyThrIleValGlyGlySerAlaGluAsp 678
QY 2701 TTGATCATCATGATGTTGTTGCTAGCAATTTACCAAGTATTGTTAGATCTAAGATAT 2760
Db 679 PheSerGluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysAspLeu 698
QY 2761 GTAGTGTGCTGTTCCATCATCATGCTTGCATAGATCCGTTTATCTCTACCCCAAA 2820
Db 699 PheIleValGluAsnThrSerHisAsnThrLeuAlaSerLeuThrLeuGlnHisArgAla 718
QY 2821 GCTTTATGTGATCTCTATTTG-----TTCGAGATGCGTTT----- 2856
Db 719 PheLeuGlyGlyLeuProMetProSerPheGlySerIleThrAspMetLeuLysAspIle 738
-----ATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCA 2901
739 ProLeuIleLeuAsnAlaGlnLeuSerThrLysAsnAspMetAspThrArg 758
2902 TATACATTTGAGAGGAGGATGTTGTTGGGATTAATACTGCTCGCTGAGAGATT 2961
759 TyrThrSerThrProGluAlaGlnGlySerThrAsnAsnSerGlyAlaLeuGluLeu 778
2962 GGAGCGGATACCGATTGTGATTACTCATCTAAGCTCTATTGATGAGTTGGCTCT 3021
779 GlyGlySerLeuAlaLeuLeuProLysGluAlaProPhePheGlnGlyThrPhePro 798
3022 TTCTGTCAGCTGAGTTTTTCTTATGCCGATCATGATCTTTTACAGAGAAGCGCATCA 3081
799 PheLeuLysPheGlnAlaValThrSerArgGlnAsnPheLysGluSerGlyAlaGlu 818
3082 GCTCGGATTCACAGAGCGGATCTCTTAATCTATCATGTTCTCTGTTGGAGTGAAGTTT 3141
819 AlaArgAlaPheAspAspGlyAspLeuValAsnCysSerIleProValGlyIleArgLeu 838
3142 GATCGATGTTCTAGTACACATCTCTAATAAATAGCTTTTATGCCGCTTATATCTGTGAT 3201
839 GluLysIleSerGluAspGluLysAsnAsnPheGluIleSerLeuAlaThrIleGlyAsp 858
3202 GCTTATCGCACCATCTCTGTTACTGAGACAACGCTCTTATCCCATCAAGAGCATGGACA 3261
859 ValThrArgLysAsnProArgSerArgThrSerLeuMetValSerGlyAlaSerTrpThr 878
3262 ACAGATGCCTTTCATTTAGCAAGACATGGAGTTGGTGTAGAGGATCTATGATGCTTCT 3321
879 SerLeuCysLysAsnLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeuThr 898
3322 CTAACAGTAATATAGATATATGCCCATGCAAGATATGATGATCGAGATGCTTCTCGA 3381
899 LeuSerProHisValGluLeuSerGlyGluAlaAlaThrGluLeuArgGlySerAlaHis 918
3382 GCGTATGTTCTAGTGCAGGAAGTAGAGTCCGGTTC 3417
919 IleThrAsnValAspCysGlyLeuArgThrSerPhe 930
RESULT 31
US-09-438-185A-448
; Sequence 448, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 448
; LENGTH: 938
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0446
; US-09-438-185A-448
Alignment Scores: 6.45e-91 Length: 938
Pred. No.: 1079.00 Matches: 303
Score: 43.2% Conservative: 160
Percent Similarity: 28.3% Mismatches: 407
Best Local Similarity: 28.3%
```



Query Match:	13.7%	Indels:	202
DB:	2	Gaps:	29
US-10-701-844-1 (1-4435) x US-09-438-185A-448 (1-938)			
QY	382	ATGCAACGCTCTTCCATAGTCTCTTCAATGATTTAGCTTATTTCTTGCTCTCT	441
DB	9	MetIyIleProLeuHibLysLeuLeuIleSerSerThrLeuVal---ThrProIleLeu	27
QY	442	TAAATGGGGGGATATCGACGAGAATC---ATGGTTCTCAAGGAATTTACGATGG	498
DB	28	LeuSerIleAlaThrTyrglyAlaAspAlaSerLeuSerProThrAspSerPheaspGly	47
QY	499	GAGAGTTAACTGATCATTTCCCTACTGTTATAGGAGATCCGAGTGGACTACTGTT	558
DB	48	AlaGlyGlySerThrPheThrProIysSerThr---AlaAspAlaAsnGlyThrAsnTyr	66
QY	559	TTTTCTGAGAGAGTTAAACATTAATAAATCTTGACAATTTCTATTCAGCTTTGCTTTA	618
DB	67	ValLeuSerGlyAsnValTyrlleAsnAspAlaGlyLysGly---ThrAlaLeuThrGly	85
QY	619	AGTTGTTTTGGAACTTATTAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACT	678
DB	86	CysCysPheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTySerPheSer	105
QY	679	TCGGAAACATACGAGCTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGA	738
DB	106	PheAsnThrValAspAlaGlySerAsnAlaGlyAlaAlaAlaSerThrThrAlaAspLys	125
QY	739	CTGTTTACTATTGAGGGTTTAAAGAAATTATCTCTTTTCCAAATTCGAATTCATTGCT	798
DB	126	AlaLeuThrPheThrGlyPheSerAsnLeuSerPhe	138
QY	799	GTAATGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACAAACA	849
DB	139	AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu	158
QY	850	TCTACACCGTCTAATAGTACTATTATTTCTTAAACAGATCTTTTGTACTCAATATGAG	909
DB	159	AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal	172
QY	910	AAGTTCTCATTTAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATCTAAGAGC	969
DB	173	---SerAsnGluAlaAsnAsnAsnGlyGlyAlaIleThrThrLysThr	187
QY	970	TTAACGGTTCAAGGAATTAGCAAGCTTTGTGTTCTTCCAAAGAAATFACTGCTCAAGCTGAT	1029
DB	188	LeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLysLeu	207
QY	1030	GGGGAGCTTGTCAAGTAGTCAACGATTTCTCTGCTATGGCTTAACGAGGCTCTATTGCC	1089
DB	208	GlyGlyAlaIleTySerSerAlaAlaAlaSerIleSerGlyAsnThrGlyGlnLeuVal	227
QY	1090	TTTGTAGCGAATGTTGCAGGACTAAGAGGGGAGGATTTGCTGTTTCAGGATGGCGAG	1149
DB	228	PheMetAsnAsn---LysGlyGluThrGlyGlyGlyAlaLeuGlyPheGlu	243
QY	1150	CAGGGAGTGTCTATCTACTTCAACAGAAAGATCCAGTAGTAAGTTTTCAGAAATACT	1209
DB	244	---AlaSerSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr	260
QY	1210	GGGTAGAGTTTGTATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTCAGGGAAC	1269
DB	261	AlaThrAsp	263
QY	1270	GTTGCTTTCTGTAATAATGGAATAACCTTTGTTCTCAACAATGTTGCTCTCTGTTTAC	1329
DB	263	---	263
QY	1330	ATTGCTGCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATACGAGAT	1389
DB	264	---AlaAla	267
QY	1390	GGAGGAGCTATCTTCTGAAGAAATGGTGGCAAGAGATCCAAATAACTCGGATCACTT	1449
DB	268	GlyGlyAlaIleTyrcysGluLys-----ThrGlyGluThrProThrLeu	282
QY	1450	TCCTTTGATGGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGTCTGGGAAAGGGGA	1509
DB	283	ThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThrGlnGlyGly	302
QY	1510	GCTATTATATGCAAAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAAATTTTAAAGGAT	1569
DB	303	AlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn	322
QY	1570	-----ATCGCTAATGATGGTGGAGGATTTATTATTAGGAGAACTCGGAGAG	1614
DB	323	ArgCysGlyAsnThrAlaAlaGlyGlyGlyAlaIleAlaIleAlaAspSerGlySer	342
QY	1615	CTCAGTTTATCTGCTGATTTATGAGAGATATTATTTTCGATGGGAATCTTTAAAGAACACC	1674
DB	343	LeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsn-----	357
QY	1675	AAAGAGATGCTGCGGATGTTAATGGGTAATCTGTCTCTCACAAGCCATTTGATGGGA	1734
DB	358	-----ThrLeuThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyLeuGly	375
QY	1735	TCGGGAGGAAATAACGACATTAAGAGCTAAACAGCGATCAGATTTCTTTAATGAT	1794
DB	376	SerSerAlaLysIleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleTyPheTyAsp	395
QY	1795	CCCATCGAGATGGCAACGGAAAT-----AACCAACCCAGCGCAG	1833
DB	396	ProIleAlaSerAsnThrThrGlyAlaSerAspValLeuThrIleAsnGlnProAspSer	415
QY	1834	TCTTCCAAACTCTTAAATAATTACGATGGTGAAGATACACAGGGATATTGTTTCTCT	1893
DB	416	AsnSerProLeu-----AspTySerGlyThrIleValPheSer	428
QY	1894	-----AATGGAAGCAGTACTTTGTAC	1914
DB	429	GlyGluLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys	448
QY	1915	CAAAATGTTACGATAGAGCAAGGAGTCTTCTTCTGTAAGGCAAAATTTATCATGTG	1974
DB	449	GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLysGlyAsnValGluLeuAspVal	468
QY	1975	AATCTCTAAGTCAGACAGGTGGAGTCTGTATATGGAAGCTGGGAGTACATGGGATTTT	2034
DB	469	AsnGlyPheThrGlnThrGluGlySerThrLeuLeu-----	480
QY	2035	GTAACCTCCACCAACCAACAGCTCTCCGCTAATCATGATTCACGCTTTCCAAT	2094
DB	481	-----MetGlnProGlyThrLysLeuLysAlaAspThrGluAlaIleSerLeuThrLys	498
QY	2095	CTGCATTTGTCTCTTCTTCTTGTAGCAACCAATGCAATTCGATTCCTCTACCAAT	2154
DB	499	LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer-----	513
QY	2155	CCTCAGCGCAAGATCTCATCTCTCCAGTCATGTTAGTACACACTCTGCT	2205
DB	514	-----IleGluThrAlaGlyAlaAsnLys	521
QY	2206	TCTGTTACAAATVAGTGGCTATCTTTTGGAGATTTGGATGATACAGCTTATGATAGG	2265
DB	522	ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyrglu---	540
QY	2266	TATGATGGCTAGGTTCTAATCAAAAAATCAATGCTGCTGAAATTTACAGTTAGGACTAAG	2325
DB	541	-----SerHisThrIleAsn-----GlnAlaPheThrGln	550
QY	2326	CCC-----CCAGCTAATGCCCATCAGATTTTACTCTTAGGG-----	2361
DB	551	ProLeuValPheThrAlaAlaSerAspIleTyIleAspAlaLeuLeu	570
QY	2362	-----AATGAGATGCTTAAGTATGCTATCAAGGAAGCTGGAAGCTTGCG	2406

Db 571 ThrSerProValGlnThrProGluProHisTyrGlyTyrGlnGlyHisTyrGluAlaThr 590  
 QY TGG---GATCCTAATACAGCAAAATTAATGGTCTCTTATCTCTGAAGCTACATGGACTAAA 2463  
 Db 591 TrpAlaAspThrSerThrAlaLysSerGly-----ThrMetThrTyrValThr 606  
 QY ACTGGGTATATCTGGCGCTGAGGAGTAGCTCTTGGTTCCAAATAGTTTATGGGGA 2523  
 Db 607 ThrGlyTyrAsnProAsnProGluArgAlaSerValProAspSerLeuTrpAla 626  
 QY TCCATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGAGCGGCTCT 2583  
 Db 627 SerPheThrAspLeuThrLeuGlnLeuMetThrSerGlnAlaAsnSerIleTyr 646  
 QY TATTCGAGGATATAGGTTTCTGGAGTTTCGAATTTCTTATCATGACCGCGATGCT 2643  
 Db 647 GlnGlnArgGlyLeuTrpAlaSerGlyThrAlaAsnPhePheHisLysAspLysSerGly 666  
 QY TTAGTCTCAGGATATCGGTATATTAGTGGGGTTTCTTAGGAGCAAACTCTAC--- 2700  
 Db 667 ThrAsnGlnAlaPheArgHisLysSerTyrGlyTyrIleValGlyGlySerAlaGluAsp 686  
 QY TTTGGATCATCGATGTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTAT 2760  
 Db 687 PheSerGluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysLeu 706  
 QY GTAGTGTGTCTTCCATCATCATGCTTGCATAGGATCGGTTTATCTATCTACCCAAAC 2820  
 Db 707 PheIleValGluAsnThrSerHisAsnTyrLeuAlaSerLeuTyrLeuGlnHisArgAla 726  
 QY GCTTTATGTGGATCTATTG-----TTCCGAGATGCGTTT----- 2856  
 Db 727 PheLeuGlyGlyLeuProMetProSerPheGlySerIleThrAspMetLeuLysAspIle 746  
 QY -----ATCCGTGTCTAGTACGGGTTTGGGAATCAGCATGATAAAACCTCA 2901  
 Db 747 ProLeuIleLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThrArg 766  
 QY TATACATTTGCAGAGAGAGCGATGTTCTGGTGGATAATACTGTCTGGCTGGAGAGATT 2961  
 Db 767 TyrThrSerTyrProGluAlaGlnGlySerTyrThrAsnAsnSerGlyAlaLeuGluLeu 786  
 QY GGAGCGGATATCCGATGTGATGATTCTCATCTAAGCTCTATTGATGATGCTGCTCT 3021  
 Db 787 GlyGlySerLeuAlaLeuTyrLeuProLysGluAlaProPhePheGlnGlyTyrPhePro 806  
 QY TTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGGAAGCGCATCAA 3081  
 Db 807 PheLeuLysPheGlnAlaValTyrSerArgGlnGlnAsnPhelGlySerGlyAlaGlu 826  
 QY GCTCGGCGCATTCAGAGCGGACATCTCCAAATCTATCAGTCTCTGTTGGAGTGAAGTTT 3141  
 Db 827 AlaArgAlaPheAspAspGlyAspLeuValAsnCysSerIleProValGlyIleArgLeu 846  
 QY GATCGATGTTCTAGTACATCTCATATAATATAGCTTTATGGCGGCTTATCTGTGAT 3201  
 Db 847 GluLysIleSerGluAspGluLysAsnAsnPheGluIleSerLeuAlaTyrIleGlyAsp 866  
 QY GCTTATCGCACCATCTCTGTGATCGACAAACCTCTCTATCCCATCAAGAGACATGGACA 3261  
 Db 867 ValTyrArgLysAsnProArgSerArgThrSerLeuMetValSerGlyAlaSerTrpThr 886  
 QY ACAGATGCTCTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGTATGCTTCT 3321  
 Db 887 SerLeuCysLysAsnLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeuThr 906  
 QY CTAACAGTATATAGAGTATATGGCCATGGAAGATATGAGTATCGAGTATCGATCTTCGA 3381  
 Db 907 LeuSerProHisValGluLeuSerGlyGluAlaAlaTyrGluLeuArgGlySerAlaHis 926  
 QY GGTATGTTGATGTCAGCAAGTAGTCCGCTTC 3417

Db 927 IleTyrAsnValAspCysGlyLeuArgTyrSerPhe 938  
 RESULT 32  
 US-09-438-185A-447  
 ; Sequence 447, Application US/09438185A  
 ; Patent No. 6822071  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stephens, Richard  
 ; APPLICANT: Mitchell, Wayne  
 ; APPLICANT: Kalman, Sue  
 ; APPLICANT: Davis, Ronald  
 ; APPLICANT: The Regents of the University of California  
 ; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence  
 ; FILE REFERENCE: 018941-000411US  
 ; CURRENT APPLICATION NUMBER: US/09/438,185A  
 ; CURRENT FILING DATE: 2002-03-13  
 ; PRIOR APPLICATION NUMBER: US 60/108,279  
 ; PRIOR FILING DATE: 1998-11-12  
 ; PRIOR APPLICATION NUMBER: US 60/128,606  
 ; PRIOR FILING DATE: 1999-04-08  
 ; NUMBER OF SEQ ID NOS: 1074  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 447  
 ; LENGTH: 947  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia pneumoniae  
 ; FEATURE:  
 ; OTHER INFORMATION: CPH0445  
 US-09-438-185A-447  
 Alignment Scores:  
 Pred. No.: 2,19e-87 Length: 947  
 Score: 1041.00 Matches: 310  
 Percent Similarity: 44.1% Conservative: 162  
 Best Local Similarity: 29.0% Mismatches: 406  
 Query Match: 13.2% Indels: 192  
 DB: 31 Gaps: 31  
 US-10-701-844-1 (1-4435) x US-09-438-185A-447 (1-947)  
 QY 382 ATGCAACGCTTTCCATTAAGTCTTTCTTCAATGATCTAGCTATTATTCTTGCTGCTCT 441  
 Db 12 MetLysSerSerValSerTrpLeuPhePheSerSerIleProLeuPheSerSerLeuSer 31  
 QY 442 TTAATATGGGGGGATATGCACGAGAAATCATGGTT-----CCTCAAGAAATTTACGAT 495  
 Db 32 Ile-----ValAlaAlaGluValThrLeuAspSerSerAsnAsnSerTyrAsp 47  
 QY 496 GGGGAGACGTTAACTGTATCTATTCCTCTATCTATTAGGAGATCCGAGTGGGACTACT 555  
 Db 48 GlySerAsnGlyThrThrPheThrValPheSerThrThrAspAlaAlaAlaGlyThrThr 67  
 QY 556 GTTTTTCGCGAGGAGATTACATTAAATAATCTTGACAATCTTATTCGAGCTTTGCTCT 615  
 Db 68 TyrSerLeuLeuSerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAla 87  
 QY 616 TTAAGTGTGTTGGGAACCTTATTGGAGCTTTTACTGTTTATAGGAGAGAGACTCGTTG 675  
 Db 88 SerGlyCysPheLeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeu 107  
 QY 676 ACTTTTCGAGAACATACCGACT---TCTACAAATGGGGCAGCTCTTAAGTAATAGCGCTGCT 732  
 Db 108 LysPheAlaPheIleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAla 127  
 QY 733 GATGAGCTGTTTACTATTAGGGGTTTAAAGAAATATCTTTTCCAAATTCGATTCATTATTA 792  
 Db 128 AspLysAsnLeuLeuPheAsnAspPheSerArgLeuSerIleIleSerCysProSerLeu 147  
 QY 793 CTTCGCGTACTGCTGCTGTCACGACTAATAGGGTAGCCAGACTCCGAGCAGCAACATCT 852  
 Db 148 LeuLeu-----SerProThr----- 152  
 QY 853 ACACCGCTTAATGGTACTATTATTCTTAAACAGATCTTTTGTGTACTCAATATGAGAAG 912

Db 153 GlyGlnCysAlaLeuLysSerValGlyAsnLeuSerLeuThrGlyAsnSerGln 170  
Qy 913 TTCTCATCTATAGTAATTTAGTCTCTCGAGATGGGAGGAGCTATAGATGCTAAAGAGCTTA 972  
Db 171 IleIlePheThrGlnAsnPheSerSerAspGlyGlyValIleAsnThrLysAsnPhe 190  
Qy 973 ACGGTTCAAGGAATTAAGCAAGCTTTGTCTCTCCAGAAATATCTGCTCAAGCTGATGGG 1032  
Db 191 LeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAla 206  
Qy 1033 GGAGCTTGCTCAAGTAGTACCAGTTCTCTGCTATGGCTAAGCAGGCTCTATTGCTCTTT 1092  
Db 206 206  
Qy 1093 GTAGCGAATGTTGAGGAGTAAGAGGGGAGGATGCTGCTTTCAGGATGGGCGAGCAG 1152  
Db 207 PheThrGlyLysGlnGlyValValTyrAla 217  
Qy 1153 GGAGTGTCTATCTATCTCAACAGAGATCCA---GTAGTAAGTTTTCAGAAATACT 1209  
Db 218 ---ThrGlyThrIleThrIleGluAsnSerProGlyIleValSerPheSerGlnAsnLeu 236  
Qy 1210 CGGTTAGAGTTGATGGCAACGTAGCCGAGTAGGAGGAGGATTTACTCTCAGCGAAC 1269  
Db 237 Ala-----LysGlySer-----GlyGlyAlaLeuTyrSerThrAspAsn 249  
Qy 1270 GTTGCTTCTCGAATAATAGCAAAACCTTTGTTCTCAACAATGTTGCTTCTCTCTTTTAC 1329  
Db 250 CysSerIleThrAspAsnPheGlnValIlePheAspGlyAsnSerAla-----Tyr 266  
Qy 1330 ATTGCTGCTAAGCAACCAAGTCGACAGGCTCTTAATACGATTAATAATACGAGAT 1389  
Db 267 GluAlaAlaGln-----AlaGln 272  
Qy 1390 CGAGAGCTATCTCTCTAAGATGGTGGCAAGCAGGATCCAATACTCTGGATCATT 1449  
Db 273 GlyGlyAlaIleCysCys-----ThrThrAspLysThrVal 285  
Qy 1450 TCCTTTGATGAGAGGAGTAGTATTTCTTAGTACAAATGTAGCTGCTGGGAAAGGGGA 1509  
Db 286 ThrLeuThrGlyAsnLysAsnLeuSerPheThrAsnAsnThrAlaLeuThrTyrGlyGly 305  
Qy 1510 GCTATTATGCAAAAGCTTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAGGAAT 1569  
Db 306 AlaIleSerGlyLeuLysValSerIleSerAlaGlyGlyProThrLeuPheGlnSerAsn 325  
Qy 1570 ATGCTAATGAT-----GGTGGAGCGATTATTAGGAGAACTCGA 1611  
Db 326 IleSerGlySerSerAlaGlyGlnGlyGlyGlyAlaIleAsnIleAlaSerAlaGly 345  
Qy 1612 GAGCTCAGTTTATCTGCTGATATGAGATATATTTTCGATGGGAATCTTAAAGAAC 1671  
Db 346 GluLeuAlaLeuSerAlaThrSerGlyAspIleThrPheAsnAsnAsn----- 361  
Qy 1672 GCCAAGAGAATGCTGCCGATTTAATGCGTAACCTGTGCTCTCAAGCCATTTTCGATG 1731  
Db 362 -----GlnValThrAsnGlySerThrSerThrArgAsnAlaIleAsnIle 376  
Qy 1732 GGATCGGAGGGAATAACGACATTAAGCATTAAGCAGGAGGATCTCTCTTAAT 1791  
Db 377 IleAspThrAlaLysValThrSerIleArgAlaAlaThrGlyGlnSerIleTyrPheTyr 396  
Qy 1792 GATCCATCTCAGATGCAAAACGGAATAACACGACGAGCTCTTCCAAACTTCTAAAA 1851  
Db 397 AspProIleThr-----AsnProGlyThrAlaAlaSerThrAspThrLeuAsn 412  
Qy 1852 ATTAACGATGGTAAGGA-----TACACAGGGAATATTTT 1890  
Db 413 LeuAsnLeuAlaAspAlaAsnSerGluIleGluTyrGlyGlyAlaIleValPheSerGly 432  
Qy 1891 -----GCTAATGGAAGCAGTACTTTGTACCAA 1917

Db 433 GluLysLeuSerProThrGluLysAlaIleAlaAlaAsnValThrSerThrIleArgGln 452  
Qy 1918 AATGTTACGATAGAGCAAGAGGATTTGCTTCTGTAAGAGGCAAAATATATCATGTGAAT 1977  
Db 453 ProAlaValLeuAlaArgGlyAspLeuValLeuArgAspGlyValThrValThrPheLys 472  
Qy 1978 TCTCTAAGTCAGACAGCTGGAGT---CTGTATATGGAAGCTGGAGTACATGGGATTTT 2034  
Db 473 AspLeuThrGlnSerProGlySerArgIleLeuMetAspGlyGlyThrThrLeuSer--- 491  
Qy 2035 GTAACTCCACAACACCAACACAGCTCTCCCTCAATCACTGATCAGCTTTTCCCAAT 2094  
Db 492 -----AlaLysGluAlaAsnLeuSerLeuAsnGly 501  
Qy 2095 CTGCAATTTGCTCTTTCTTTTGTAGCAACAATGCGAGTTTACGATCTCTCCCTACCAAT 2154  
Db 502 LeuAlaValAsnLeuSerSerLeuAspGlyThrAsnLysAla----- 515  
Qy 2155 CCTCAGCGCAAGATTCTCATCTCCGATCTATTGGTAGCACCACTGCTGGTTCTGTTACA 2214  
Db 516 -----AlaLeuLysThrGluAlaAlaAspLysAsnIleSer 527  
Qy 2215 ATTAGTGGGCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGTATGATGTTGG 2274  
Db 528 LeuSerGlyThrIleAlaLeuIleAspThrGluGlySerPheTyrGluAsnHisAsn--- 546  
Qy 2275 CTAGGTTCTAATAAATAATCAATGCTGCTGAAATTTACAGTTAGGAGCTAAGCCCCAGCT 2334  
Db 547 LeuLysSerAlaSerThrTyrProLeuLeuGluLeu-----ThrThrAlaGlyAla 563  
Qy 2335 AAT-----GCCCATCAGATTGACTCTAGGGAATGAGATGCTTAAG 2376  
Db 564 AsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGlnGluProGluThrHis 583  
Qy 2377 TATGCTATCAAGAGCTGGAAGCTTGGTGGATCTCTAATACAGCAAAATAATGGTCT 2436  
Db 584 TyrGlyTyrGlnGlyAsnTyrGlnLeuSerTyrPalaAsnAlaThrSerSerLys----- 601  
Qy 2437 TATACTCTGAAGCTTACATGCACTAAACTGGGTATATCTGGGCCCTGAGGAGTAGCT 2496  
Db 602 ---IleGlySerIleAsnTyrThrArgThrGlyTyrIleProSerProGluArgLysSer 620  
Qy 2497 TCTTTGGTCCCAATAGTTTATGGGATCCATTTTATAGATATACGATTCGCGATTCAGCA 2556  
Db 621 AsnLeuProLeuAsnSerLeuTyrPalaAsnIleAspIleArgSerIleAsnGlnLeu 640  
Qy 2557 ATTCAAGCAAGTGGATGGGCGCTCTTATTGTCAGGATTTATGGGTTTCTGGAGTTTGG 2616  
Db 641 IleGluThrLysSerSerGlyGluProPheGluArgGluLeuTyrLeuSerGlyIleAla 660  
Qy 2617 AATTCTTCTATCATGACGGGATGCTTTAGTACGGATATCGGTATATATGTCGGGT 2676  
Db 661 AsnPhePheTyrArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly 680  
Qy 2677 TATTCTTCTAGGACCAACTCTCTACTTT---GGATCATCGATGTTTGGTCTAGCATTTACC 2733  
Db 681 TyrAlaLeuGlyIleThrAlaThrThrProAlaGluAspGlnLeuThrPheAlaPheCys 700  
Qy 2734 GAAGTATTGTTGATAGTAAAGATTATGTAGTGTCTGTTCCCAATCATCATCTGTCATA 2793  
Db 701 GlnLeuPheAlaArgAspAsnHisIleThrGlyLysAsnHisGlyAspThrTyrGly 720  
Qy 2794 GGATCCGTTTATCTATCTACCCCAACAAGCTTTA-----TGTGGATCTCTATTGTTGGA 2847  
Db 721 AlaSerLeuTyrPheHisHisThrGluGlyLeuPheAspIleAlaAsnPheLeuTyrGly 740  
Qy 2848 GATCGGTTT----- 2856  
Db 741 LysAlaThrArgAlaProTyrValLeuSerGluIleSerGlnIleIleProLeuSerPhe 760  
Qy 2857 ---ATCCGCTGCTAGCTAGCGGTTTGGGAATCAGCATATGAAACCTCATATATGCA 2913  
Db 761 AspAlaLysPheSerTyrLeuHisThrAspAsnHisMetLysThrTyrTyrThr----- 778



QY 1447 GTTCTCTTGATGAGAGGAGTAGTTTCTTCTTAGTAGCAATGATGCTGCGGAAGGG 1506  
DDB : : : : :  
278 LeuThrLeuSerAspAenglyGluLeuAsnPhelileGlyAenThrAlalleThrSerGly 297  
QY 1507 GGAGCTATTATGCAAAAGCTCTCGGTCTCTAAGCTGTCGCTCTGACATTTTAAAG 1566  
DDB : : : : :  
298 GlyAlalleThrAspAsnLeuValLeuSerSerGlyGlyProThrLeuPheLeuAsn 317  
QY 1567 AATATCGCTAATGAT-----CGTCGAGCGATTTATTATGAGGAGATCTGGA 1611  
DDB : : : : :  
318 AsnSerAlalleAspThrAlaAlaProLeuGlyGlyAlalleAlalleAlaAspSerGly 337  
QY 1612 GAGCTCAGTTATCTGCTGATATATGAGATATTATTTTCGATGGGAATCTTAAAGACA 1671  
DDB : : : : :  
338 SerLeuSerLeuSerAlaLeuGlyGlyAspileThrPheGluGlyAsn-----Thr 354  
QY 1672 GCCAAGAGATGTCGCGGATGTTAATGCGCTAACTGTGCTCTACAAGCCATTTCGATG 1731  
DDB : : : : :  
355 ValVallyysGlyAlaSer-----SerSerGlnThrThrThrArgAsnSerileAsnile 372  
QY 1732 GGA---TCGGGAGGAAATAACGACATTAAGAGCTAAAGCAGGCGATCAGATTTCTTT 1788  
DDB : : : : :  
373 GlyAenThrAenAlaLysileValGlnLeuArgAlaSerGlnGlyAsnThrileThrPhe 392  
QY 1789 AATGATCCATCGAGATGCCAAACGGAATAACACGAGCGGCGAGTCTTCCAAACTTCTA 1848  
DDB : : : : :  
393 TyrAspProileThr-----ThrSerileThrAlaAlaLeuSerAspAlaLeu 408  
QY 1849 AAAATTAACGATGGTAA-----GGATACACAGGGGATTTGTTTTGCT 1893  
DDB : : : : :  
409 AsnLeuAsnGlyProAspLeuAlaGlyAsnProAlaTyrglnGlyThrileValPheSer 428  
QY 1894 -----AATGGAAGCAGTACTTTGTAC 1914  
DDB : : : : :  
429 GlyGluLysLeuSerGluAlaGluAlaAlaLeuAspAsnLeuLysSerThrileGln 448  
QY 1915 CAAATGTTACGATAGACAGAGGAGGATTTCTTCGTGAAAAGCGCAAAATATACATG 1974  
DDB : : : : :  
449 GlnProLeuThrLeuAlaGlyGlyGlnLeuSerLeuLysSerGlyValThrLeuValAla 468  
QY 1975 AATCTCTAAGTCACAGAGTGGAGT---CTGTATATGGAAGCTGGAGTACATGGAT 2031  
DDB : : : : :  
469 LysSerPheSerGlnSerProGlySerThrLeuLeuMetAspAlaGlyThrThrLeuGlu 488  
QY 2032 TTTGTAATCCCAACACCACACAGCTCTCGCGCTAATCATGTTGATCAGCTTTCC 2091  
DDB : : : : :  
489 ThrAlaAspGly-----IleThrileAsn 496  
QY 2092 AATCTGCATTGCTCTTCTTCTTTGTTAGCAAAATGCAATGCAATCTCTCTACC 2151  
DDB : : : : :  
497 AsnLeuValLeuAsnValAspSerLeu----- 505  
QY 2152 AATCTCCAGCCAGATTTCTCATCTCTGAGTCATGTTGATGCAACATCTGCTGT---TCT 2208  
DDB : : : : :  
506 -----LysGlnThrLysLysAlaThrLeuLysAlaThrGlnAlaSerGlnThr 521  
QY 2209 GTTACATTAAGTGGGCTATCTTTTGGAGTATGGATGATACAGCTTATGATAGTAT 2268  
DDB : : : : :  
522 ValThrLeuSerGlySerLeuValAspProSerGlyAsnValTyrgluAspVal 541  
QY 2269 GATTCGCTAGGTTCTAATCAAAAATCAATGTCCTGMAATACAGTTAGGAGCTAAGCCC 2328  
DDB : : : : :  
542 SerTrp-----AsnAsnProGlnVal---PheSerCysLeuThrLeuThrAlaAspAsp 558  
QY 2329 CAGCTAATGCCCCA---TCAGATTGACT-----CTAGGGGAATGAGTGCCTAAG 2376  
DDB : : : : :  
559 ProAlaAsnileHisileThrAspLeuAlaAlaAspProLeuGluLysAsnProileHis 578  
QY 2377 TATGCTATCAGGAGCTGGAGCTTGGTGGATCTTAATACAGCAAAATATGCTCT 2436  
DDB : : : : :  
579 TrpGlyTyrglnGlyAsnTrpAlaLeuSerTrpGlnGluAspThrAlaThrLysSerLys 598

QY 2437 TATACTCTGAAGCTACATGGACTAAACTGGGTATTAATCTGGCCTGAGCGAGTAGCT 2496  
DDB : : : : :  
599 ---AlaAlaThrLeuThrTrpThrLysThrGlyTyrAsnProAsnProGluArgGly 617  
QY 2497 TCTTTGGTTCCAAATAGTTATGGGATCCATTTAGATATACCATTCGCGCATTCAGCA 2556  
DDB : : : : :  
618 ThrLeuValAlaLeuThrLeuTrpGlySerPheValAspValArgSerIleGlnGlnLeu 637  
QY 2557 ATTCAACAAAGTGGATGGCGCTCTATTGTGTCAGGATATGGTTCTCGAGTTTCG 2616  
DDB : : : : :  
638 ValAlaThrLysValArgGlnSerGlnGlnThrArgGlyileTrpCysGluGlyIleSer 657  
QY 2617 AATTTCTTCTATCATGACCGCGATGCTTTAGTTCAGGATATCGTATATATAGTGGGGT 2676  
DDB : : : : :  
658 AsnPhePheHisLysAspSerThrLysileAsnLysGlyPheArgHisileSerAlaGly 677  
QY 2677 TATTCCTTAGGCAAACTCTCTACTTTGGATCA---TCGATGTTTGGTCTAGCATTTACC 2733  
DDB : : : : :  
678 TyrValValGlyAlaThrThrThrLeuAlaSerAspAsnLeuIleThrAlaAlaPheCys 697  
QY 2734 GAAGTATTGCTAGATCTAAAGATTATGTAGTGTGCTTCCATCATCATCTGTTGCATA 2793  
DDB : : : : :  
698 GlnLeuPheGlyLysAspArgAspHisPheileAsnLysAsnArgAlaSerAlaTyraAla 717  
QY 2794 GGATCCGTTTATCTATCTACCCCAAGCTTTATGTGGA-----TCCTATTGG 2841  
DDB : : : : :  
718 AlaSerLeuHisLeuGlnHisLeuAlaThrLeuSerSerProSerLeuLeuArgTyrlLeu 737  
QY 2842 TTCGGA-----GATCGTTTATCCGTGCTAGCTACCGG 2874  
DDB : : : : :  
738 ProGlySerGluSerGluGlnProValLeuPheAspAlaGlnile-----SerTyrlle 755  
QY 2875 TTTGGGAATCAGCATATGAAAACCTCATATATTTGCAGAGAGAGCGATGTTGTTGG 2934  
DDB : : : : :  
756 TyrSerLysAsnThrMetLysThrTyrrTyrrGlnAlaProLysGlyGluSerSerTrp 775  
QY 2935 GATTAATTAATCTCTGCTGGGAGATGCGGCGGATTCACGATTTGATTTGATTTCCACT 2994  
DDB : : : : :  
776 TyrAsnAspGlyCysAlaLeuGluLeuAlaSerSerLeuProHisThrAlaLeuSerHis 795  
QY 2995 AAGCTCTATTGAAATGAGTTCGCTTCTGCAAGCTGAGTTTCTTATTCGCGCATCAT 3054  
DDB : : : : :  
796 GluGlyLeuPheHisAlaTyrrPheProPheileLysValGluAlaSerTyrlleHisGln 815  
QY 3055 GAATCTTTTACAGAGAGGCGCATCAA---GCTCGGCAATCAAGAGCGGACATCTCTA 3111  
DDB : : : : :  
816 AspSerPheLysGluArgAsnThrThrLeuValArgSerPheAspSerGlyAspLeuile 835  
QY 3112 AATCTATCAGTTCTGTTGGAGTGAGTTTATGATGATGTTCTTAGTACATCTCTAATAA 3171  
DDB : : : : :  
836 AsnValSerValProileGlyileThrPheGluArgPheSerArgAsnGluArgAlaSer 855  
QY 3172 TATAGCTTTATGGCGCTTATATCTGATGCTTATCGCACCATCTCTGTTACTGAGACA 3231  
DDB : : : : :  
856 TyrGluAlaThrValileTyrrValAlaAspValTyrrArgLysAsnProAspCysThrThr 875  
QY 3232 ACGTCTCTATCCCATCAAGAGACATGGAACAAGATGCTTTCTATTTAGCAAGACATGGA 3291  
DDB : : : : :  
876 AlaLeuLeuileAsnAsnThrSerTrpLysThrThrGlyThrAsnLeuSerArgGlnAla 895  
QY 3292 GTTGTGGTTAGGATCTATGATGCTTCTTAACAGTAATATAGAAGTATATGCGCAT 3351  
DDB : : : : :  
896 GlyileGlyArgAlaGlyilePheTyrrAlaPheSerProAsnLeuGluValThrSerAsn 915  
QY 3352 GGAAGATATGATCTCGAGATGCTTCTCGAGCTATGTTTGGTGCAGGAAGTAGATGC 3411  
DDB : : : : :  
916 LeuSerMetGluileArgGlySerSerArgSerTyrrAsnAlaAspLeuGlyLysPhe 935  
QY 3412 CGGTTT 3417  
DDB : : : : :  
936 GlnPhe 937

US-09-198-452A-472  
; Sequence 472, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Grifais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
; TITLE OF INVENTION: and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 472  
; LENGTH: 927  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 1...927  
; OTHER INFORMATION: Xaa=unknown or other  
US-09-198-452A-472

Alignment Scores:  
Pred. No.: 4,48e-83 Length: 927  
Score: 994.50 Matches: 308  
Percent Similarity: 43.8% Conservative: 157  
Best Local Similarity: 29.0% Mismatches: 413  
Query Match: 12.6% Indels: 183  
DB: 2 Gaps: 38

US-10-701-844-1 (1-4435) x US-09-198-452A-472 (1-927)

QY	382	ATGCAACGCTCTTCCATAGTCTTTCTTCAATGATCTAGCTTATCTTGCTGCTCT	441
DB	1	MetLysSerSerLeuHisTrpPheLeuIleSerSerSerLeuAlaLeuPro---LeuSer	19
QY	442	TTAAATGGGGGGATATGACGAGAAATCATGGTT-----CCTCAAGGAATTTAC	492
DB	20	LeuAsnPheSerAlaPheAlaAlaValValGluIleAsnLeuGlyProThrAsnSerPhe	39
QY	493	GATGGGAG---ACGTAACTGTATCATTTCCCTTACTGTTATAGGAGTCCGAGTGGG	549
DB	40	SerGlyProGlyThrTyThr-----ProProAlaGlnThrThrAsnAlaAspGly	56
QY	550	ACTACTGTTTTTCTGCGAGAGATTAACATTAATAAAATCTTGACAAATCTATTGAGCT	609
DB	57	ThrIleTyAsnLeuThrGlyAspValSerIleThrAsnAla---GlySerProThrAla	75
QY	610	TTGCCCTTAAGTTGTTTGGAACTTATTAGGAGTTTACTGTTTATAGGAGGACAC	669
DB	76	LeuThrAlaSerCysPheLysGluThrThrGlyAsnLeuSerPheGlnGlyHisGlyTyr	95
QY	670	TCGTTGACTTTCGAGAACATACGACTCTTCAAAATGGGCGAGCTCTAAGTAAATAGCGCT	729
DB	96	GlnPheLeuGlnAsnIleAspAlaGlyAlaAsn---CysThrPheThrAsnThrAla	114
QY	730	GCTGATGAGCTGTACTATATGAGGTTTAAAGAAATTAATCTTTTCCAAATTCGAATCA	789
DB	115	AlaAsnLysLeuLeuSerPheSerGlyPheSerTyThrLeuSerLeu-----	129
QY	790	TTACTTCCCGTACTGCTGTCGCAACGACTAATAAGGTAGCCAGACTCCGACGACACACA	849
DB	130	-----IleGlnThrThrAsn	134
QY	850	TCTACACCGCTAATGCTACTATTATTCTAAACAGATCTTTTGTGTACTCAATAATGAG	909
DB	135	AlaThrThrGlyThrGlyAlaIleLysSerThrGlyAlaCysSerIleGlnSerAsnTy	154
QY	910	AAGTTCTCTATATAGTATTTAGTCTCTGAGATGGGGGAGCTATAGATGCTAAGACG	969
DB	155	SerCysTyThrPheGlyGlnAsnPheSerAsnAspAsnGlyAlaLeuGlnGlySerSer	174
QY	970	TTAACGGTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAGAAAATACTGCTCAAGCTGAT	1029

DB	175	IleSerLeu---SerLeuAsnProAsnLeuThrPheAlaLysAsnLysAlaThrGlnLys	193
QY	1030	GGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCTAACGAGGCTCTATTGCC	1089
DB	194	GlyGlyAlaLeu-----TyrSerThr	200
QY	1090	TTTGTAGCGAATGTTGAGGAGTAGAGGGGGAGGATGCTGCTGTTTCAGGATGGCAG	1149
DB	201	-----GlyGlyIleThr-----	204
QY	1150	CAGGAGTGTCAATCATCTACTTCAACAGAAGATCCAGTAGTAAGTTTTCAGAAATACT	1209
DB	205	-----IleAsnAsnThrLeuAsnSer-----AlaSerPheSerGluAsnThr	218
QY	1210	GGGTAGAGTTTGTAGGGAACGTAGCCGAGTAGGAGGAGGATTTACTTCTACGGGAAC	1269
DB	219	Ala-----AlaAsnAsnGlyGlyAlaIleTyThrGluAlaSer	231
QY	1270	GTTGCTTCTCGAATAATAGAAACCTTG---TTTCTCAACAATGTTGCTTCTCTCTGTT	1326
DB	232	---SerPheIleSerSerAsnLysAlaIleSerPheIleAsnAsn-----	245
QY	1327	TACATTGCTCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATTACGGA	1386
DB	246	SerValThrAlaThrSerAlaThr-----	253
QY	1387	GATGGAGAGCTATCTTCTGTGAAGATGCTGGCGCAAGCAGGATCCAAATCACTCTGATCA	1446
DB	254	---GlyGlyAlaIleTyCys-----SerSerThrSerAlaProLysProVal	268
QY	1447	GTTTCTTGTAGGAGGAGTGTCTTCTTAGTACCAATGATGCTCTCGGAAAGG	1506
DB	269	LeuThrLeuSerAspAsnGlyGluLeuAsnPheIleGlyAsnThrAlaIleThrSerGly	288
QY	1507	GGAGCTATTATGCCAAAAGCTCTCGTGTGCTAACTGCGCCCTGTACAAATTTTAAAGG	1566
DB	289	GlyAlaIleTyThrAspAsnLeuValLeuSerSerGlyGlyProThrLeuPheLysAsn	308
QY	1567	ATATCGCTAATGAT-----GCTGGAGCGATTTATTAGGAGAACTCTGGA	1611
DB	309	AsnSerAlaIleAspThrAlaAlaProLeuGlyGlyAlaIleAlaIleAlaAspSerGly	328
QY	1612	GAGCTCAGTTTATCTGCTATTATGAGATATTTTCATGGGGAATCTTAAAAACA	1671
DB	329	SerLeuSerLeuSerAlaLeuGlyAspIleThrPheGluGlyAsn-----Thr	345
QY	1672	GCCAAAGGAATGCTCCCGATGTTAATGGCGTAATGCTCTCTCAACGCCATTTCTGATG	1731
DB	346	ValValLysGlyAlaSer-----SerSerGlnThrThrThrArgAsnSerIleAsnIle	363
QY	1732	GGA---TCGGGAGGGAATAATACGACATTAAGAGCTAAAGCAGGCGATCAGATCTCTTT	1788
DB	364	GlyAsnThrAsnAlaLysIleValGlnLeuArgAlaSerGlnGlyAsnThrIleTyPhe	383
QY	1789	AATGATCCCATCGAGTAGCCAAACGAAATAACCCAGCCAGCGAGCTTCCAACTCTTA	1848
DB	384	TyrAspProIleThr-----ThrSerIleThrAlaAlaLeuSerAspAlaLeu	399
QY	1849	AAAAATTAACCATGGTGA-----GGATACACAGGGGATATTGTTTGTCT	1893
DB	400	AsnLeuAsnGlyProAspLeuAlaGlyAsnProAlaTyGlnGlyThrIleValPheSer	419
QY	1894	-----AATGAGACGACTCTTGTGATC	1914
DB	420	GlyGluLysLeuSerGluAlaGluAlaAlaGluAlaAspAsnLeuLysSerThrIleGln	439
QY	1915	CAAAATGTTACGATAGCAAGGAGGATTTCTTCTGTAAGAAAGCAAAATATCAGTG	1974
DB	440	GlnProLeuThrLeuAlaGlyGlnLeuSerLeuLysSerGlyValThrLeuValAla	459
QY	1975	AATTTCTTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGGGAGTACATGGAT	2031





326 SerPheThrSerAsnThrThrAlaAlaAsnGlyGlyAlaIleTyrAlaThrLysCysThr 345  
Qy  
475 GTTCTCAAGGAATTACGATGGGAGACGTTAACTGTATCATCTCCCTATCTGTTATA 534  
Db  
346 Leu-----AspGlyAsnThr-----ThrLeuThrPheAspGlnAsnThrAla 359  
Qy  
535 GGAGATCCGAGTGGGACTACTGTTTCTCGCAGGAGAGTTAAACATTAATAAAATCTTCAC 594  
Db  
360 ThrAlaGlyCysGlyAlaIleTyrThrGluThrGlu----- 372  
Qy  
595 AATTCTATTGCAGCTTTCCTTTAAAGTTGTTTGGGAACCTATTAGGAGGTTTACTGTT 654  
Qy  
373 -----AspPheSerLeu 376  
Db  
655 TTAGGAGAGACACTCGTTGACTTTCGAGACATACCGACTTCTACAAATGGGCGAGCT 714  
Qy  
377 LysGlySerThrGlyThrValThrPheSerThrAsnThrAlaLysThrGlyGlyAlaLeu 396  
Db  
715 CTAAGT-----AATAGCGCTGCTGATGGA-----CTGTTTACTATTGAGGT 756  
Qy  
397 TyrSerLysGlyAsnSerSerLeuThrGlyAsnThrAsnLeuLeuPheSer-----Gly 414  
Db  
757 TTTAAGAATATTCCTTTTCCAAATTCGAATTCATTACTTGCCTGCTGCTGCTGCAACG 816  
Qy  
415 AsnLysAlaThrGlyProSerAsn-----SerSer 424  
Db  
817 ACTAATAAGGTAGCCAGACTCCGACGACACATCTACACCGTCTAATGGTACTATTAT 876  
Qy  
425 AlaAsnGlnGluGlyCysGlyAlaIleLeuAlaPheIleAspSerGlySerValSer 444  
Db  
877 TCTAATAACAGATCTTTGTTACTCAATATGAGAAGTTCTCAATCTATAGTAATTTAGTC 936  
Qy  
445 AspLysThrGlyLeuSerIleAlaAsnAsnGlnGluValSerLeuThrSerAsnAlaAla 464  
Db  
937 TCTGAGATGGGGAGCTATAGATCTTAAGACTTAAACGGTTCAAGGAATTACGAGCTT 996  
Qy  
465 ThrValSerGlyAlaIleTyrAlaThrLysCysThrLeuThrGlyAsnGlySerLeu 484  
Db  
997 TGTGCTTCCAGAAATACTCTCAAGCTGATGGGGAGCTTGTCAGTAGTACC----- 1053  
Qy  
485 ---ThrPheAspGlyAsnThrAlaGlyThrSerGlyGlyAlaIleTyrThrGluThrGlu 503  
Db  
1054 AGTTTCTGCTATGCTAAGCAGGCTCTATTGCTTGTAGCGAATGTCAGGAGTA 1113  
Qy  
504 AspPheThrLeuThrGlySerThrGlyThrValThrPheSerThrAsnThrAlaLysThr 523  
Db  
1114 AGAGGGGGAGGATT----- 1128  
Qy  
524 ---GlyGlyAlaLeuTyrSerLysGlyAsnAsnSerLeuSerGlyAsnThrAsnLeuLeu 542  
Db  
1129 -----GCTGCTGTTCCAGATGGGCGACG 1152  
Qy  
543 PheSerGlyAsnLysAlaThrGlyProSerAsnSerSerAlaAsnGlnGluGlyCysGly 562  
Db  
1153 CGAGTG-----TCATCATCTACTTCAACA----- 1176  
Qy  
563 GlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysGlyLeuTrpIle 582  
Db  
1177 ---GAAGATCCAGTAGTAGTTTTCAGAAATATCTGCG----- 1212  
Qy  
583 GluAspAsnGluAsnValSerLeuSerGlyAsnThrAlaThrValSerGlyAlaIle 602  
Db  
1213 -----GTAGATTTCATGGGACGTAGCC 1236  
Qy  
603 TyrAlaThrLysCysAlaLeuHisGlyAsnThrThrLeuThrPheAspGlyAsnThrAla 622  
Db  
1237 CGA---GTAGGAGGAGGATTAC-----TCCTAC 1263  
Qy  
623 GluThrAlaGlyGlyAlaIleTyrThrGluThrGluAspPheThrLeuThrGlySerThr 642  
Db  
1264 GGGACGTTGCTTCTCTGTAATAT-----GGA 1290  
Qy  
643 GlyThrValThrPheSerThrAsnThrAlaLysThrAlaGlyAlaLeuHisThrLysGly 662  
Db

1291 AAAACCTTGTTTCTCAACAATGTTGCTTCTCTCTGTTTACATTTGCTTAAGCAACCAACA 1350  
Qy  
663 AsnThrSerPheThrLysAsnLysAla-----LeuValPheSerGlyAsnSerAla 679  
Db  
1351 AGTGCAGAGCTTCTTAATACGAGTAATAATTAACGAGATGAGAGGACTATCTTCTGT--- 1407  
Qy  
680 ThrAlaThrAlaThrThrThrAspGlnGluGlyCysGlyGlyAlaIleLeuCysAsn 699  
Db  
1408 -----AAGAATGGTGGCAAGCA 1425  
Qy  
700 IleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAsnGluSerLeuSer 719  
Db  
1426 GGATCCCAATAACTCTGGATCAGTTTTC-----TTT 1455  
Qy  
720 PheIleAsnAsnThrAlaLysArgSerGlyGlyGlyIleTyrAlaProLysCysValIle 739  
Db  
1456 GATGCAGAGGAGTAGTATTTCTTTAGTAGCAATAGTAGCTGCGGAAAGGGGAGCTATT 1515  
Qy  
740 SerGlySerGluSerIleAsnPheAspGlyAsnThrAlaGluThrSerGlyGlyAlaIle 759  
Db  
1516 TATGCCAAAAGCTCTCGTTGCTAACTGTGCGCTGTACAATTTTAAAGGAATATCGCT 1575  
Qy  
760 TyrSerLysAsnLeuSerIleThrAlaAsnGlyProValSerPheThrAsnAsnSerGly 779  
Db  
1576 AATGATGCTGGAGCGATTATTTAGGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTAT 1635  
Qy  
780 GlyLysGlyGlyAlaIleTyrIleAlaAspSerGlyGluLeuSerLeuGluAlaIleAsp 799  
Db  
1636 GGAGATATTTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCCGATGTT 1695  
Qy  
800 GlyAspIleThrPheSerGlyAsn-----ArgAlaThrGluGlyThrSerThrPro 816  
Db  
1696 AATGCGCTAACTGCTCTCACAGCCATTTCCATGGGATCGGAGGAGAAATAACGACA 1755  
Qy  
817 Asn-----SerIleHisLeuGlyAlaGlyAlaLysIleThrLys 829  
Db  
1756 TTAAGAGCTTAAGCAGGCGATCAGATTCTCTTTAATATGATCCCATCGAGATG--- 1806  
Qy  
830 LeuAlaAlaIleProGlyHisThrIleTyrPheTyrAspProIleThrMetGluAlaPro 849  
Db  
1807 GCAAAACGGA----- 1815  
Qy  
850 AlaSerGlyGlyThrIleGluGluLeuValIleAsnProValValLysAlaIleValPro 869  
Db  
1816 -----AATAACCGCCAGCGAGCTCTCCAACTTCTAAATTAACGATGCT 1863  
Qy  
870 ProProGlnProLysAsnGlyProIleAlaSerValProValProValAlaProAla 889  
Db  
1864 GAAGGATACACAGGCGATATGTTTT----- 1890  
Qy  
890 AsnProAsnThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAspAlaSer 909  
Db  
1891 -----GCTAATGGAAGCAGTACTTTGTACCAAAATGTTTACGATAGACAGCAAGAGATT 1944  
Qy  
910 IleProAlaAsnThrThrThrIleLeuAsnGlnLysIleAsnLeuAlaGlyGlyAsnVal 929  
Db  
1945 GTTCTTGTGAAAGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTTGGAGCT--- 2001  
Qy  
930 ValLeuLysGluGlyAlaThrLeuGlnValTyrSerPheThrGlnGlnProAspSerThr 949  
Db  
2002 CTGTATATGGAAGTGGGAGTACATGGGATTTTGTAACTCCACACCAACCAACAGCCT 2061  
Qy  
950 ValPheMetAspAlaGlyThrThrLeuGluThrThrThrAsn----- 964  
Db  
2062 CTGCGCGTAAATCAGTTGATCAGCTTTCCAAATCTGCATTTGCTCTTTCTTTCTTTGTTA 2121  
Qy  
965 ---AsnThrAspGlySerIleAspLeuLysAsnLeuSerValAsnLeuAspAlaLeuAsp 983  
Db  
2122 GCAAACATGCAATGCAATGCAATCTCTACCAATCTCCAGCGCAAGATTCTCATCTCGCA 2181  
Qy  
984 GlyLysArgMetIleThr----- 989  
Db



1204 AATACTGCGGTAGAGTTTGTATGCGAAGCTAGCCGAGTAGGAGGAGGATTTACTCTAC 1263  
77 -----GlyAlaAlaIlePheThrGlu 83  
1264 GCGAAGCTGTCTTCTGTAATATGGAATAACCTTGTCTCAACAATGTCTCTCTCT 1323  
84 AlaSerValThrIleSerAsnAlaLysValSerPheIleAspAsn ----- 99  
1324 GTTTACATTGCTGTAAGCAACCAAGTGTACAGGCTCTTAATACGAGTAATAATYAC 1383  
100 -----LysValThrGlyAlaSerSerThrThrGlyAspMet 112  
1384 CGAGATGGAGGAGCTATCTTCTGTAAAGATGTGCGCAAGCAGGATCAATAACTCTGGA 1443  
113 -----SerGlyGlyAlaIleCysAlaThrLys -----ThrSerThrAspThr 126  
1444 TCAGTTTCTCTTGTATGAGGAGGAGTGTATTTCTTTAGTAGCAATGTAGCTGTGGAAA 1503  
127 LysValThrLeuThrGlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrAla 146  
1504 GGGGAGCTATTATGCCAAAGCTCTCGGTGTCTAACTGGCCCTGTACAAATTTTA 1563  
147 GlyAlaIleThrValLysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSer 166  
1564 AGCAATATCGCTAAT -----GATGTGGAGCGATTTATTAGGAGATCT 1608  
167 ArgAsnSerValAsnGlyThrAlaProLysGlyGlyAlaIleAlaIleAlaIleGluAspSer 186  
1609 GGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGA 1668  
187 GlyGluLeuSerLeuSerAlaAspSerGlyAspIleValPheLeuGlyAsn ----- 203  
1669 ACAGCCAAAGAGAACTGCTGCCGATGTAAATGGCTTAATGCTGTCTCACAGCCATTCG 1728  
204 ThrValThrSerThrThrProGlyThrAsn -----ArgSerSerIleAsp 218  
1729 ATGGATCGGGAGGAAATAACGACATTAAGACTAAAGCAGGATCAGATTCCTTT 1788  
219 LeuGlyThrSerAlaLysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleThrPhe 238  
1789 AATGATCCCATCCAGATGCGCAAGGAAATAACAGCCAGCGAGTCTTCCAACTTCTA 1848  
239 TyrAspProIleThrThrGlySerSerThr -----ValThrAspValLeu 254  
1849 AATAATTAACGATGGTGAAGGA -----TACAGCGGATATGTTTGTCT 1893  
255 LysValAsnGluThrProAlaAspSerAlaLeuGlnThrGlyAsnIleIlePheThr 274  
1894 -----AATGAGCAGTACTTTGTAC 1914  
275 GlyGluLysLeuSerGluThrGluAlaAlaAspSerLysAsnLeuThrSerLysLeuLeu 294  
1915 CAAAATGTTACGATAGACGAAGGATGTTCTTCGTAAGGCAAGGCAAAATTAATCAGTG 1974  
295 GlnProValThrLeuSerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThr 314  
1975 AATTCTCTAAGTCAGACAGGTGGAGT -----CTGTATATGAAGCTGGAGTACATGGAT 2031  
315 GluAlaPheThrGlnGlnAlaAspSerArgLeuGluMetAspValGlyThrThrLeuGlu 334  
2032 TTTGTAACTCCACACCACCAACAGCCCTCTGCGCTAATCAGTTGATCAGCTTTC 2091  
335 -----ProAlaAspThrSer -----ThrIleAsn 342  
2092 AATCTGCATTTGCTCTTCTTCTTCTTGTAGCAACAACTAGCAGTTAGCAATCTCTTACC 2151  
343 AsnLeuValIleAsnIleSerSerIle ----- 351  
2152 AATCTCCAGCGCAAGATTTCTACCTGCAATCTGTTAGC -----ACAACCTGCTGTCT 2208  
352 -----AspGlyAlaLysLysAlaLysIleGluThrLysAlaThrSerLysAsn 367  
2209 GTTACAATTAGTGGCCCTATCTTTTGTGAGGATTTGGATGATACAGCTTATAGGTAT 2268

368 LeuThrLeuSerGlyThrIleThrLeuLeuAspProThrGlyThrPheThrGluAsnHis 387  
2269 GATTGCGCTAGTCTTAATCAAAAATCAATGTCCTGAAATACAGTTA -----GGGACTAAG 2325  
388 Ser -----LeuArgAsnProGlnSerTyrAspIleLeuGluLeuLysAlaSerGlyThrVal 406  
2326 CCCCAGCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCCTAAGTATGCTAT 2385  
407 ThrSerThrAlaValThrProAspProIleMetGlyGluLysPhe -----HisTyrGlyTyr 425  
2386 CAAGGAAGCTGG -----AAGCTTGGTGGATCTCTTAATACAGCAATAATAGTCTTATCT 2442  
426 GlnGlyThrTrpGlyProIleValThrGlyThrGlyAlaSerThrThrAlaThrPhe ----- 444  
2443 CTGAAAGCTACACTGACTGAACTGAGTATATCTGGCCCTGAGGAGTAGTCTTCTTTG 2502  
445 -----AsnTrpThrLysThrGlyTyrIleProAsnProGluArgIleGlySerLeu 461  
2503 GTTCAAATAGTTTATGGGATCCATTTTATAGATATACGATCTGCGCATTCAGCAATCAA 2562  
462 ValProAsnSerLeuTrpAsnAlaPheIleAspIleSerSerLeuHisTyrLeuMetGlu 481  
2563 GCAAGTGTGCTGGCGCTCTTATGTCGAGGATTAATGGTTTCTCGAGTTTCGATTC 2622  
482 ThrAlaAsnGluGlyLeuGlnGlyAspArgAlaPheTrpCysAlaGlyLeuSerAsnPhe 501  
2623 TTTCTATCATGACCGCATCTTTAGTCAAGGATATCGTATATATAGTGGGTATTC 2682  
502 PheHisLysAspSerThrLysThrArgArgGlyPheArgHisLeuSerGlyGlyTyrVal 521  
2683 TTAGGAGCAACTCTCTACTTTTGGATCA -----TCGATGTTTGGTCTAGCATTTACCGAAGTA 2739  
522 IleGlyGlyAsnLeuHisThrCysSerAspLysIleLeuSerAlaAlaPheCysGlnLeu 541  
2740 TTTGTPAGATCTAAAGATTAATGATGTGTCTTCAATCATCATGCTTGCATAGATTC 2799  
542 PheGlyArgAspArgAspTyrPheValAlaLysAsnGlnGlyThrValTyrGlyGlyThr 561  
2800 GTTATCTATCTACCAACAGCTTTATGCGATCTTATGTTTCTCGGAGATGCTTTATC 2859  
562 LeuTyr-TyrGlnHisAsnGluThrTyrIleSerLeuProCysLysLeuArgProCys 581  
2860 CGTGCT -----AGCTACCGGTT 2876  
581 rLeuSerTyrValProThrGluIleProValLeuPheSerGlyAsnLeuSerTyrThrHi 601  
2877 TGGGAATCAGCATATGAAAACCTCATATACATTTGAGAGGAGGAGGATGTTCTGGGA 2936  
601 sThrAspAsnAspLeuLysThrLysTyrThrTyrProThrValLysGlySerTrpG1 621  
2937 TAATAACTGTCTGCTGGAGAGATTCGAGCGGATTTACCGATTTGTGATTTACTCCATCTAA 2996  
621 yAsnAspSerPheAlaLeuGluPheGlyGlyArgAlaProIleCysLeuAspGluSerAl 641  
2997 GCTCTATTTGAATGAGTGTGCTCTTTGCGCAAGCTGAGTCTTCTATCCCATCATGA 3056  
641 aLeuPhe -----GluGlnTyrMetProPheMetLysLeuGlnPheValTyrAlaHisGlnG1 660  
3057 ATCTTTTACAGAGAAAGCGCATCAAGCTCGGCAATTCAGAGCGGACATCTCTTAATCT 3116  
660 uGlyPheLysGluGlnGlyThrGluAlaArgGluPheGlySerSerArgLeuValAsnLe 680  
3117 ATCAGTTCTGTGGAGTGAAGTTTGTATCGATGTTCTTAGTACATCTTAATAATATAG 3176  
680 uAlaLeuProIleGlyIleArgPheAspLysGluSerAspCysGlnAspAlaThrTyrAs 700  
3177 CTTTATGCGGCTTATCTGTGCTTATCGACCATCTTATCGACCATCTCTGGTACTGAGACAACGCT 3236  
700 nLeuThrLeuGlyTyrThrValAspLeuValArgSerAsnProAspCysThrThrLe 720  
3237 CCTATCCCATCAAGAGACATGGCAACAGATGCTCTTTCATTTTAGCAAGACATGGAGTTGT 3296

```
Db 720 wArgIleSerGlyAapSerTprLysThrPheGlyThrAsnLeuAlaArgGlnAlaLeuVa 740
Qy 3297 GGTAGAGATCTATGATGCTTCTTAACAAAGTAATATAGAGATATATGGCCATGGAG 3356
Db 740 lLeuArgAlaGlyAanHisPheCysPheAsnSerAsnPheGluAlaPheSerGlnPheSe 760
Qy 3357 ATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGTAGTCAGGAAGTAGTCCGGTT 3416
Db 760 rPheGluLeuArgGlySerSerArgAsnTyrAsnValAspLeuGlyAlaLysTyrGlnPh 780
Qy 3417 C 3417
Db 780 e 780

RESULT 37
US-09-438-185A-6
Sequence 6, Application US/09438185A
Patent No. 6822071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-000411US
CURRENT APPLICATION NUMBER: US/09/438,185A
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 932
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
FEATURE:
OTHER INFORMATION: Cpn0005
US-09-438-185A-6

Alignment Scores:
Pred. No.: 4,456-81 Length: 932
Score: 973.00 Matches: 306
Percent Similarity: 41.5% Conservative: 148
Best Local Similarity: 28.0% Mismatches: 386
Query Match: 12.3% Indels: 253
DB: 2 Gaps: 35

US-10-701-844-1 (1-4435) x US-09-438-185A-6 (1-932)
Qy 375 TTGTGTGATGCAACGCTCTTCCATAAGTTCTTCTTCAAT-----GAT 419
Db 15 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAap 31
Qy 420 TCTAGCTATTCTGCTCTCTTAAATGGGGGGGATATGAGCAGAAATCATGGTCC 479
Db 32 ThrSerLeu-SerAlaThrThrIleSer-----LeuThrPr 43
Qy 480 TCAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATCTGTTATAGGAGA 539
Db 43 oGluAapSerPheHisGlyAapSerGlnAsnAlaGluArgSerTyrAsnVal----- 60
Qy 540 TCCGAGTGGGACTACTGTTTTCGCGAGGAGTTAAACATTAATAAATCTTGACAATTC 599
Db 61 -GlnAlaGlyAapValTyrSerLeuThrGlyAapValSerIleSerAsnValAspAsnSe 80
Qy 600 TATTGACAGCTTTGCTTTAAAGTTGTTTGGGAACCTATTATAGGAGTTTCTGTTTAGG 659
Db 80 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGl 98
Qy 660 GAGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACA-----AATGGGGCAGC 713
```

```
Db 98 yAsnHisGlyLeuTyrPheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVa 118
Qy 714 TCTAAGTAATAGCGCTGCTGATGACTGTTTACTATATTGAGGGTTTTAAAGAAATATCCTT 773
Db 118 lLeuCysCysGlnAapProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 138
Qy 774 TTCCAATTGCNAATTCATTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCA 833
Db 138 e-----IleGl 140
Qy 834 GACTCCGACGACAACTCTACACCGTCTAATGGTACTATTATTCTATAAACACAGATCTTTT 893
Db 140 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTyrSerLysAsnAlaLeuWe 158
Qy 894 GTTACTCAATAATAGAGATTCTCATTTATAGTAATTTAGTCTCTGAGAGATGGGGAGC 953
Db 158 tLeuLeuAsnAsnTyrValValArgPheGluGlnAsnGlnSerLysThrLysGlyGlyAl 178
Qy 954 TATAGATGCTAAGAGCTTAACGGTTCAAGGNAATTAGCAAGCTTTGTCTCTCCAGAAAA 1013
Db 178 aIleSerGlyAlaAsnValThrIleValGlyAsnTyrAspSerValSerPheTyrGlnAs 198
Qy 1014 TACTGCTCAAGCTCATGGGGAGCTTGTCAAGTAGTCACAGTTTCTCTGCTATGGCTAA 1073
Db 198 nAlaAlaThrPheGlyGly-----AlaIleHisE 208
Qy 1074 CGAGGCTCTATTCCTTTTGTAGGAATGTTGCGAGAGTAAGAGGGGAGGATTGCTGC 1133
Db 208 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 222
Qy 1134 TGTTCAGGATGGGAGGAGGAGTGTCTATCTACTTCAACAGAGATCCAGTAGTAAG 1193
Db 222 ----- 222
Qy 1194 TTTTCCAGAAATACTGCGTAGAGTTTGTATGGAAACGTAGCCCGAGTAGGAGGAGGAT 1253
Db 223 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 236
Qy 1254 TTACTCTCAGGAAACGTTGCTTTCTGTAATTAATGAAACCTTTGTTTCTCAACAATGT 1313
Db 236 uTyr-SerAspGlyAapIleAapGlnAsnAlaTyrValLeuPheArgGluAsnGl 256
Qy 1314 TGCTTCTCTGTTTACATTGCTGTCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAG 1373
Db 256 uAlaLeuThrThrAlaIle----- 262
Qy 1374 TAATAATTACCGAGATGGAGGAGCTATCTTCTGTAAGAATGTTGCGCAAGCAGGATCCA 1433
Db 263 -----GlyLysGlyAlaValCysCys-----LeuPr 272
Qy 1434 TAACCTGATGATCA-----GTTTCTTGTATGAGAGGAGTAGTTT 1475
Db 272 oThr-SerGlySerThrProValProIleValThrPheSerAapAsnLysGlnLeuVa 292
Qy 1476 CTTTAGTAGCAATGTAGCTGCTGGGAAGGGGAGCTATTATTATGCCAAAAGCTCTCGGT 1535
Db 292 lPheGluArgAsnHisSerIleMetGlyGlyAlaIleTyrAlaArgLysLeuSerIl 312
Qy 1536 TGCTAACTGTGCGCTGTACAAATTTTAAAGCAATATCGCT-----AATGA 1580
Db 312 eSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTyrAlaAsnSerGlnAsnLe 332
Qy 1581 TGGTGGAGCGATTATTATTAGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGA 1640
Db 332 uGlyGlyAlaIleAlaIleAapThrGlyGlyGluIleSerLeuSerAlaGluLysGlyTh 352
Qy 1641 TATTATTTTCGATGGGAATCTTAAAGAACAGCAAGAGATGTCGCGATGTTAATGG 1700
Db 352 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGl 367
Qy 1701 CGTAACTGTGCTCTCACAAGGCCATTTCGATGGGATCGGGAGGAGAAATAACGACATTAG 1760
Db 1760
```

Db 367 ylleHisLeuLeuGlnAsnAla-----LysPheLeuLeuLysLeuGl 380  
QY 1761 AGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCATCGAGATGGCAACGGAATAA 1820  
Db 380 nAlaArgAsnGlyTyrSerIleGluPheTyrAspProIle-----Th 394  
QY 1821 CCAGCCAGGCGAGTCTTCCAAACTCTTAAAAATTAAAC-----GATGGTGAAGGATA 1871  
Db 394 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 414  
QY 1872 CACAGGGGATATGTTTTGCT-----AATGGAG 1901  
Db 414 rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAsnAspProArgAspPheTy 434  
QY 1902 CAGTACTTTGTACCAAAATGTTACGATAGACAGAGGATGTTCTTCGTGAAAGGC 1961  
Db 434 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyrLeuValIleGlyGluGlyAl 454  
QY 1962 AAAATATCAGTGAATCTCTAAGTCAGACAGTGGAGT---CTGTATATGGAAGCTGG 2018  
Db 454 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuGl 474  
QY 2019 GAGTACATGGGATTTTGTAACTCCACACCACACAGCCCTCTGCGCTAATCAGTT 2078  
Db 474 yThr-----LysLeuIleAlaSerLysGluAs 483  
QY 2079 GATCAGCTTTCGAATCTGATTTGCTCTTCTTCTTCTTTGTTAGCAAAATGCAGTTAC 2138  
Db 483 pileAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 496  
QY 2139 GAATCTCTTACCAATCTCCAGCGCAAGATTCATCTCGCAGTCATTGTTAGCACAA 2198  
Db 497 -----SerSerSerThrAlaAlaValIleLysAlaAsnTh 509  
QY 2199 TGCTGCT-----TCTGTTACAATTAGTGGCGCTATCTTT----- 2232  
Db 509 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 529  
QY 2233 -TTTGAGGATTTGGATGATACAGCTTATGATGATGATGATGCTAGGTCTTAATCAAAA 2291  
Db 529 atyTyrGluAspLeuArg-----MetArgAsnSerGlnTh 540  
QY 2292 AATCAATGCTCTGAATACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATT 2351  
Db 540 rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlyGlySerValThrVa 556  
QY 2352 GACTTAGGGAATGAGATG-----CCTAAGTATGGCTATCAAGGAAGCTGGAAGCT 2402  
Db 556 lThrAlaGlyAspPheLeuProValSerProHisTyrGlyPheGlnGlyAsnTrpLysLe 576  
QY 2403 TGCGTGGGATCCTAATACAGCAAAATATGCTCTTATCTCTGAAAAGCTACATGGACTAA 2462  
Db 576 uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-----PheTrpAspLy 592  
QY 2463 AACTGGGTATATCTCGGGCTGAGCGAGTAGCTCTTTGTTCCAAATAGTTTATGGGG 2522  
Db 592 sIleAsnTyrLysProArgProGluLysGluGlyAsnLeuValProAsnIleLeuTrpGl 612  
QY 2523 ATCCATTTTATAGATATAGCTGGCATTCCAGCAATTCAGCAAGTGTGATGGCGCTC 2582  
Db 612 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuGl 632  
QY 2583 TTAFTTTCGAGGATTTATGGTTCCTGAGATTTTCAATTTCTTCATCATCACCGCGATGC 2642  
Db 632 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPhePheHisValSerAlaSerGl 652  
QY 2643 TTTAGTCCAGGATATCGGTATATAGTGGGGTTATCTTACGACCAAACTCTACTT 2702  
Db 652 uAspAsnIleArgTyrArgHisAsnSerGlyTyrValLeuSerValAsnAsnGluIl 672  
QY 2703 TGGATCATCGATGTTT---GGTCTAGCATTTTACCAAGATATTTGGTAGATCTAAAGATTA 2759  
Db 672 eThrProLysHisTyrThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 692

QY 2760 TGTAAGTGTGCTTCCCAATCATGCTTGCATAGGATCGTGTATCTATCTTACCCAACA 2819  
Db 692 rAlaValSerAsnAsnGluTyrArg-----MetTyrLeu----- 703  
QY 2820 AGCTTTATGTGATCTCTATTTGTTCT-----GGAGATGCGTTTATCCG 2861  
Db 704 -----GlySerTyrLeuTyrGlnTyrThrThrSerLeuGlyAsnIlePheArgTy 720  
QY 2862 TGCTAGC----- 2868  
Db 720 rAlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnPr 740  
QY 2869 -----TACGGGTTTGGGAATCAGCATATGATAAAACCTC 2900  
Db 740 oLeuMetIlePheHisPheLeuCysAlaTyrGlyHisAlaThrAsnAspMetLysThrAs 760  
QY 2901 ATATCATTTTGCAGAGGAGGATGTTGCTGGGATATACTACTGCTGGCTGGAGAGAT 2960  
Db 760 pTyrAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCy 780  
QY 2961 TGGAGCGGATTTACGATTTGATTTACTCCATCTAAGCTCTATTGAAATGAGTTGCGTCC 3020  
Db 780 sGlyGlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr 800  
QY 3021 TTTCTGCTCAAGCTGAGTTTCTTATGCTCCGATCATGAATCTTTTACAGAGGAGCGCATCA 3080  
Db 800 oPheMetLysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheLysGluThrThrAlaAs 820  
QY 3081 AGCTCGGCGATTCAGAGCGGACATCTCTCTAAATCTATCAGTTCTCTGTTGGAGTGAATT 3140  
Db 820 pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 840  
QY 3141 TGATCGATGTTCTAGTACACATCTCTAATAAATATAGCTTTATGGCGGCTTATATCTGTGA 3200  
Db 840 eGluLysLeuAlaLeuSerGlnAspValLeuTyrAspPheSerPheSerTyrIleProAs 860  
QY 3201 TGCTTATCCACCATCTCTGCTAGTACAGACAAACGCTCTATCCATCAAGACAGATGGAC 3260  
Db 860 pilePheArgLysAspProSerCysGluAlaLeuValIleSerGlyAspSerTrpLe 880  
QY 3261 AACAGATGCTTTTCATTTAGCAAGACATGGAGTTGGTTAGAGGATCTATGTATGCTTC 3320  
Db 880 uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyrHI 900  
QY 3321 TCTAACAGTATATAGATATATGCGCCATCGAAGATATAGTATCGAGATGCTTCTCG 3380  
Db 900 sPheAsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 920  
QY 3381 AGGCTATGTTTGGTGCAGGAGTAGAGTCCGGTTC 3417  
Db 920 gAsnTyrAsnIleAsnCysGlySerLysPheArgPhe 932

## RESULT 38

US-09-198-452A-15  
; Sequence 15 Application US/09198452A  
; Patent No. 655294  
; GENERAL INFORMATION:  
; APPLICANT: Grifais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 15  
; LENGTH: 922  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 1...922

OTHER INFORMATION: Xaa=unknown or other  
US-09-198-452A-15

Alignment Scores:  
Pred. No.: 3.03e-80 Length: 922  
Score: 964.00 Matches: 304  
Percent Similarity: 41.4% Conservative: 149  
Best Local Similarity: 27.8% Mismatches: 387  
Query Match: 12.2% Indels: 253  
DB: 2 Gaps: 35

US-10-701-844-1 (1-4435) x US-09-198-452A-15 (1-922)

```

QY 375 TTGTGTGATGCAAGCGTCTTCCATAAGTCTTTCTTTCAAT-----GAT 419
DB 5 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 21
QY 420 TCTAGCTTATCTTCTGCTCTTTAAATGGGGGGGATATGCAGCAGAAATCATGTTCC 479
DB 22 ThrSerLeu-SerAlaThrThrIleSer-----LeuthrPr 33
QY 480 TCAAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATACTGTTTATAGGAGA 539
DB 33 oGluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTyrAsnVal----- 50
QY 540 TCCGAGTGGGACTACTGTTTTTTCGAGGAGAGTTAACTAAATAATCTTGACAATTC 599
DB 51 -GlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 70
QY 600 TATTGACGCTTGCCTTTAAAGTTGTTTGGGAACCTATTAGGAGTCTTTCTGTTTATAGG 659
DB 70 r-----AlaLeuAsnLysAlaCysPhe***ValThrSerGlySerValThrPheAlaGl 88
QY 660 GAGAGGACACTGTTGACTTTCGAGAACATACGAGCTTCTACA-----AATGGGGCAGC 713
DB 88 yAsnHisHisGly***TyrPheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVala 108
QY 714 TCTAAGTAATAGCGTCTGTGATGAGTGTCTTACTATGAGGTTTAAAGAAATATCTTT 773
DB 108 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLysSerPh 128
QY 774 TTCCAATTGCAATTCATTACTTTCGCTGCTGCTGCTCAACGACTAATAAGGTTAGCCA 833
DB 128 e-----AsnGl 130
QY 834 GACTCCGACGACACATCTACACCGTCTAATGGTACTATTATTCTTAAACAGACTTTT 893
DB 130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTyrSerLysAsnAlaLeuMe 148
QY 894 GTTACTCAATAATAGAAATCTCTATTCTATAGTAATTTAGTCTCTGGAGATGGGGGAGC 953
DB 148 tLeuLeuAsnAsnTyrValValArgPheGluGlnAsnGlnSerLysThrLysGlyGlyAl 168
QY 954 TATAGATGCTAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAGAAA 1013
DB 168 aIleSerGlyAlaAsnValThrIleValGlyAsnTyrAspSerValSerPheTyrGlnAs 188
QY 1014 TACTGCTCAAGCTGATGGGGGAGCTGTCAAGTAGTGCACCACTTCTCTGCTATGCTAA 1073
DB 188 nAlaAlaThrPheGlyGly-----AlaIleHisE 198
QY 1074 CGAGGCTCTATTGCTTTTGTAGCGAATGTTCCAGAGGTAAGCGGGGAGGAGATTGCTGC 1133
DB 198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 212
QY 1134 TGTTCAGGATGGGACGAGGAGTGTCACTCATCTTCTTCAACAGAGATCCAGTAGTAG 1193
DB 212 ----- 212
QY 1194 TTTTTCAGAAATACTCGCGTAGAGTTTGTATGGGAACGTAGCCCGAGTAGGAGGAGAT 1253
DB 213 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 226

```

```

QY 1254 TTACTCTACGGGAACGTTGCTTCTCTGAATAATGGAACCACTTGTCTTCTCAACAATGT 1313
DB 226 uTyrSerAspGlyAspIleAspIleAspGlnAsnAlaTyrValLeuPheArgGluAsnGl 246
QY 1314 TGCTTCTCTCTGTTTACATTTGCTGTAAAGCAACCAAGTGCAGGCGCTTCTAATACGAG 1373
DB 246 uAlaLeuThrThrAlaIle----- 252
QY 1374 TAATAATTTACGAGATGGAGGAGCTATCTTCTGTGAAGAATGGTGGCAAGCAGGATCCAA 1433
DB 253 -----GlyLysGlyGlyAlaValCysCys-----LeuPr 262
QY 1434 TAACTCTCGATCA-----GTTCTCTTTCATGGAGGAGGAGTAGTTT 1475
DB 262 oThrSerGlySerSerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 282
QY 1476 CTTTAGTAGACATGTAGCTGCTGGGAAAGGGGAGCTATTTATGTCACAAAAGCTCTCGGT 1535
DB 282 lPheGluArgAsnHisSerIleMetGlyGlyAlaIleTyrAlaArgLysLeuSerIl 302
QY 1536 TGCTAACTGTGGCCTGTACAAATTTTAAAGGAATATCGCT-----AATGA 1580
DB 302 eSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTyrAlaAsnSerGlnAsnLe 322
QY 1581 TGGTGGAGCGATTTATTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTATGAGA 1640
DB 322 uGlyGlyAlaIleAlaIleAspThrGlyGlyIleSerLeuSerAlaGluLysGlyTh 342
QY 1641 TATTATTTCGATCGGAATCTTAAAGAACAGCAGCAAGAGAAATGCTGCCGATGTTAAAGG 1700
DB 342 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGl 357
QY 1701 CGTAACTGTGTCTCACAGCCATTTTCATGGGATCGGAGCGGAAATAACGACATTAAG 1760
DB 357 yIleHisLeuLeuGlnAsnAla-----LysPheLeuLysLeuGl 370
QY 1761 AGCTAAAGCAGGATCATGATCTCTTTAATGATCCCATCGAGATGGCAACCGGAATAA 1820
DB 370 nAlaArgAsnGlyTyrSerIleGluPheTyrAspProIle-----Th 384
QY 1821 CCAGCCAGCGCAGTCTTCCAAACTTCTAAATAATTAAAC-----GATGGTGAAGGATA 1871
DB 384 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404
QY 1872 CACAGGGGATATGTTTGTCT-----AATGGAAG 1901
DB 404 rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAsnAspProArgAspPheLy 424
QY 1902 CAGTACTTTGTACCAAAATGTTACGATAGACAGGAGGATTTGTTCTGTTGAAAAGGC 1961
DB 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyrLeuValIleLysGluGlyAl 444
QY 1962 AAAATTATCAGTGAATCTCTAAGTCACAGAGTGGGAGT---CTGTATATGGAAGCTGG 2018
DB 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuGl 464
QY 2019 GAGTACATGGGATTTTGTAACTCCACACACCAACAGCCTCTCGCGCTTAATCAGTT 2078
DB 464 yThr-----LysLeuIleAlaSerLysGluAs 473
QY 2079 GATCAGCTTTCCAATCTGCATTTGTCTCTTCTTTCTTTGTAGCAACAATGCGATTAC 2138
DB 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 486
QY 2139 GAATCTCTCAACATCTCCAGCGCAAGATTCTCATCTCGCAGTCAATTGGTAGCACAA 2198
DB 487 -----SerSerSerThrAlaAlaValIleLysAlaAsnTh 499
QY 2199 TGCTGGT-----TCTGTTACAAATTAGTGGCGCTATCTTT----- 2232
DB 499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519
QY 2233 -TTTGAGGATTTGGATGATACAGCTTATGATGATGATTGCTAGGTTCTTAATCAAAA 2291

```







```

QY 2797 TCCGTTTATCTATCTACCAACAGCTTTATGTGGA-----TCCTATTGTTCCGAGAT 2850
DB 637 SerLeuArgLeuGlnHisAspAlaSerLeuTyrSerValValSerIleLeuLeuGlyGlu 656
QY 2851 GCGTTTATCCGT----- 2862
DB 657 GlyGlyLeuArgGluIleLeuLeuProTyrValSerLeuThrLeuProCysSerPheTyr 676
QY 2863 -----GCTAGCTACGGTTTGGGAATCAGCATATGAAACCTCATAT----- 2904
DB 677 GlyGlnLeuSerTyrGlyHisThrAspHisArgMetLeuThrGluSerLeuProPro 696
QY 2905 -----ACATTTGCAGAGAGAGCGATGTTCTGCGGTAATAAATCTGCTGCGTCGAGAG 2958
DB 697 ProProThrLeuSerThrAsp---HisThrSerTrpGlyTyrValTrpAlaGlyGlu 715
QY 2959 ATTGAGCGGGATTACCGATTGTTGATTAATCCATCTTAAGCTCTATTGTAATGAGTTGCGT 3018
DB 716 LeuGlyThrArgValAlaValGlnAlaValGlnAlaValGlnAlaValGlnAlaValGln 735
QY 3019 CTTTTCGTCAGAGCTGAGTTTCTTATGCGCATCATGATCTTTTACAGAGGAGCGCAT 3078
DB 736 ProPheValLysValGlnAlaValTyrAlaArgGlnAspSerPheValGluLeuGlyAla 755
QY 3079 CAAGCTCGGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTTGGAGTGAAG 3138
DB 756 IleSerArgAspPheSerAspSerHisLeuTyrAsnLeuAlaIleProLeuGlyIleLys 775
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGCGGCTTATCTCTGT 3198
DB 776 LeuGluLysArgPheAla-----GluGlnTyrTyrHisValValAlaMetTyrSerPro 793
QY 3199 GATGCTTATCGACCATCTCTGTACTGAGACACCGCTCTATCCATCCATCAAGAGCATGG 3258
DB 794 AspValCysArgSerAsnProLysCysThrThrThrLeuLeuSerAsnGlnGlySerTrp 813
QY 3259 ACAACAGATCGCTTTTCATTTAGCAGACATGCGGATGTTGTTAGAGGATCTATGATGCT 3318
DB 814 LysThrLysGlySerAsnLeuAlaArgGlnAlaGlyIleValGlnAlaSerGlyPheArg 833
QY 3319 TCTCTAACAAATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTTCT 3378
DB 834 SerLeuGlyAlaAlaGluLeuPheGlyAsnPheGlyPheGluTrpArgGlySerSer 853
QY 3379 CGAGGCTATGTTGATGTCAGGAAGTAGATCGCGGTTCC 3417
DB 854 ArgSerTyrAsnValAspAlaGlySerLysIleLysPhe 866

RESULT 40
US-09-198-452A-466
; Sequence 466, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifflis, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 466
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-466

Alignment Scores:
Pred. No.: 1,44e-67 Length: 1132
Score: 828.00 Matches: 291
Percent Similarity: 37.1% Conservative: 117
Best Local Similarity: 26.5% Mismatches: 327

```

```

Query Match: 10.5% Indels: 365
DB: 2 Gaps: 35
US-10-701-844-1 (1-4435) x US-09-198-452A-466 (1-1132)
QY 493 GATGGGAGAGCTTAACCTGATATCATTTCCCTATATCTGTTATAGGAGATCCGAGTGGACT 552
DB 36 AsnGlySerGlySerAlaAlaPheThrAlaLysGluThrSerAlaSerGlyThr 55
QY 553 ACTGTTTTTCTGCAGGAGAGTTACATTAATAAATCTTGACAAATCTTGACAAATCTTATTCAGCTTGG 612
DB 56 ThrTyrThrLeuThrSerAspValSerIleThrAsnVal---SerAlaIleThrProAla 74
QY 613 CTTTAAAGTGTGTTTGGGAATTAATTAGGGAGTTTACTGTTTATAGGGAGAGACACTCG 672
DB 75 AspLysSerCysPheThrAsnThrGlyGlyAlaLeuSerPheValGlyAlaAspHisSer 94
QY 673 TTGACTTTCGAGAACATACGGACTTCTCAAAATGCGGAGCTCTAAGTAATAGCGCTGCT 732
DB 95 LeuValLeuGlnThrIleAlaLeuThrHisAspGlyAlaAlaIleAsnAsnThrAsnThr 114
QY 733 GATGACTGTTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTA 792
DB 115 -----AlaLeuSerPheSerGlyPheSerSerLeu 124
QY 793 CTT----- 795
DB 125 LeuIleAspSerAlaProAlaThrGlyThrSerGlyGlyLysGlyAlaIleCysValThr 144
QY 795 ----- 795
DB 145 AsnThrGluGlyThrAlaThrPheThrAspAsnAlaSerValThrLeuGlnLysAsn 164
QY 795 ----- 795
DB 165 ThrSerGluLysAspGlyAlaAlaValSerAlaTyrSerIleAspLeuAlaLysThrThr 184
QY 796 -----GCCGCTAGCTGCTGCT 810
DB 185 ThrAlaAlaLeuLeuAspGlnAsnThrSerThrLysAsnGlyGlyAlaLeuCysSerThr 204
QY 811 GCAACGACT-----AATAGGCTAGCCAGACT---CCGACGACCAATCTACA 855
DB 205 AlaAsnThrThrValGlnGlyAsnSerGlyThrValThrPheSerSerAsnThrAlaThr 224
QY 856 CCGTCTAATGCTACTATTATCT----- 879
DB 225 AspLysGlyGlyIleTyrSerLysGluLysAspSerThrLeuAspAlaAsnThrGly 244
QY 879 ----- 879
DB 245 ValValThrPheLysSerAsnThrAlaLysThrGlyGlyAlaTrpSerSerAspAsn 264
QY 879 ----- 879
DB 265 LeuAlaLeuThrGlyAsnThrGlnValLeuPheGlnGluAsnLysThrThrGlySerAla 284
QY 879 ----- 879
DB 285 AlaGlnAlaAsnAsnProGluGlyCysGlyAlaIleCysCysTyrLeuAlaThrAla 304
QY 880 -----AAACACAGATCTTTTGTGTTACTCAATATGAGAAGTTCTCAATCTATATAATTAA 933
DB 305 ThrAspLysThrGlyLeuAlaIleSerGlnAsnGlnGluMetSerPheThrSerAsnThr 324
QY 934 GTCTCTGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTAAGCAAG 993
DB 325 ThrThrAlaAsnGlyGlyAlaIleTyrAlaThrLysCysThrLeuAspGlyAsnThrThr 344
QY 994 CTTTGTGCTTCCAGAAATATCTGCTCAAGCTGAT---GGGGAGCTTCTCAAGTAGTC 1050
DB 345 Leu---ThrPheAspGlnAsnThrAlaThrAlaGlyCysGlyGlyAlaIleTyrThrGlu 363

```



Db 1041 LysAspTyrValValSerAspIleLysSerGlnValTyrAlaGlySer-----LeuCys 1058  
 Qy 2812 ACCCAACAGCTTTATGTGGATCCTATTGTTCGGAGATGGCTTTATCCGTGCTAGCT 2869  
 Db 1059 AlaGlnSerSer-TyrValIleProLeuHisSerSerLeuArgArgHisValLeuSer 1077

Search completed: May 13, 2006, 10:40:44  
 Job time : 424 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:20:04 ; Search time 48 seconds  
(without alignments)  
1743.079 Million cell updates/sec

Title: US-10-701-844-2  
Perfect score: 5267  
Sequence: 1 MQTSFHKFLSMILAYSCCS.....YEVDRASRGVGLSAGSRVRF 1012

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5 COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/6 COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/H COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/PCITUS COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5267	100.0	1012	2	US-09-612-402B-2
2	5267	100.0	1012	2	US-09-542-520-2
3	5116	97.1	984	2	US-09-612-402B-43
4	5106.5	97.0	1013	2	US-09-612-402B-15
5	5106.5	97.0	1013	2	US-09-612-402B-16
6	5106.5	97.0	1013	2	US-09-542-520-15
7	5106.5	97.0	1013	2	US-09-542-520-16
8	5090	96.6	1006	2	US-09-556-877-190
9	5090	96.6	1006	2	US-09-620-412C-190
10	5090	96.6	1006	2	US-09-598-419-190
11	5084	96.5	982	2	US-09-556-877-176
12	5084	96.5	982	2	US-09-620-412C-176
13	5084	96.5	982	2	US-09-598-419-176
14	3336.5	63.3	670	2	US-10-197-220-169
15	2552	48.5	505	2	US-09-612-402B-17
16	2547	48.4	505	2	US-09-542-520-17
17	2350	44.6	458	2	US-09-612-402B-36
18	2350	44.6	458	2	US-09-542-520-36
19	1735	32.9	325	2	US-09-612-402B-37
20	1735	32.9	325	2	US-09-542-520-37
21	1365.5	25.9	999	2	US-09-438-185A-455
22	1359.5	25.8	973	2	US-09-430-723-2
23	1128	21.4	928	2	US-09-428-122-2
24	1090	20.7	949	2	US-09-198-452A-478
25	1090	20.7	967	2	US-09-438-185A-453
26	1083	20.6	930	2	US-09-198-452A-470
27	1079	20.5	938	2	US-09-438-185A-448

28	1041	19.8	947	2	US-09-438-185A-447	Sequence 447, App
29	1021	19.4	937	2	US-09-438-185A-449	Sequence 449, App
30	994.5	18.9	927	2	US-09-198-452A-472	Sequence 472, App
31	985.5	18.7	1414	2	US-09-438-185A-446	Sequence 446, App
32	980.5	18.6	780	2	US-09-438-185A-17	Sequence 17, Appl
33	967	18.4	932	2	US-09-438-185A-6	Sequence 6, Appl
34	958	18.2	922	2	US-09-198-452A-15	Sequence 15, Appl
35	939.5	17.8	866	2	US-09-438-185A-15	Sequence 15, Appl
36	821.5	15.6	1132	2	US-09-198-452A-466	Sequence 466, App
37	808	15.3	634	2	US-09-438-185A-451	Sequence 451, App
38	804	15.3	643	2	US-09-198-452A-474	Sequence 474, App
39	769.5	14.6	880	2	US-09-556-877-175	Sequence 175, App
40	769.5	14.6	880	2	US-09-620-412C-175	Sequence 175, App
41	769.5	14.6	880	2	US-09-598-419-175	Sequence 175, App
42	762.5	14.5	866	2	US-09-556-877-189	Sequence 189, App
43	762.5	14.5	866	2	US-09-620-412C-189	Sequence 189, App
44	762.5	14.5	866	2	US-09-598-419-189	Sequence 189, App
45	725	13.8	597	2	US-09-198-452A-29	Sequence 29, Appl

## ALIGNMENTS

## RESULT 1

US-09-612-402B-2  
; Sequence 2, Application US/09612402B  
; Patent No. 6642023  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; Patent No. 6642023  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/09/612,402B  
; PRIOR FILING DATE: 2000-07-06  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 1012  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-09-612-402B-2

Query Match	100.0%;	Score	5267;	DB 2;	Length	1012;
Best Local Similarity	100.0%;	Pred. No.	0;			
Matches 1012;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;					
Qy	1	MQTSFHKFLSMILAYSCCSLNGGGYAAETWVPGIYDGETLTVSPPTVIGDPSTTVF	60			
Db	1	MQTSFHKFLSMILAYSCCSLNGGGYAAETWVPGIYDGETLTVSPPTVIGDPSTTVF	60			
Qy	61	SAGELTKNLDNSIAALPLSCFGLNLSFTVLRGHSITPENITRTNGAALSNSAADGL	120			
Db	61	SAGELTKNLDNSIAALPLSCFGLNLSFTVLRGHSITPENITRTNGAALSNSAADGL	120			
Qy	121	FTTEGPKELFSNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF	180			
Db	121	FTTEGPKELFSNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF	180			
Qy	181	YSNLVSGDGAIDAKSITVQGISKLVQFQENTQAADGACQVVTSTFSAMANEAPIAFVAN	240			
Db	181	YSNLVSGDGAIDAKSITVQGISKLVQFQENTQAADGACQVVTSTFSAMANEAPIAFVAN	240			
Qy	241	VAGVRGGIAAQQGQGVSSSTSTEDPVVVSFRNTAVEFDGNVARYGGIYSYGNVAF	300			
Db	241	VAGVRGGIAAQQGQGVSSSTSTEDPVVVSFRNTAVEFDGNVARYGGIYSYGNVAF	300			
Qy	301	NGKTLPLNNVAPVYIAAKQPTSGQASNTNNYGDGGAIFCKNGAAGSNSSVSF	360			
Db	301	NGKTLPLNNVAPVYIAAKQPTSGQASNTNNYGDGGAIFCKNGAAGSNSSVSF	360			

```
QY 361 EGVVFPSSNVAAGKGAAYAKKLSVANGCPVQFLRNANDGGAIYLGSGELSLADYGD 420
Db |||||
QY 361 EGVVFPSSNVAAGKGAAYAKKLSVANGCPVQFLRNANDGGAIYLGSGELSLADYGD 420
Db |||||
QY 421 IIFDGNLKRKTAKENAADVNGVTVSSQAISMGGKITTILRAKAGHQIILFNDPIEMANGNN 480
Db |||||
QY 421 IIFDGNLKRKTAKENAADVNGVTVSSQAISMGGKITTILRAKAGHQIILFNDPIEMANGNN 480
Db |||||
QY 481 QPAQSSKLLKINDGSGYTGDIIVFANGSSTLYQNVTIEQGRIVLRKAKLSVNSLSQTGGS 540
Db |||||
QY 481 QPAQSSKLLKINDGSGYTGDIIVFANGSSTLYQNVTIEQGRIVLRKAKLSVNSLSQTGGS 540
Db |||||
QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNHLSSLLANNAVTPNPPAODSHPA 600
Db |||||
QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNHLSSLLANNAVTPNPPAODSHPA 600
Db |||||
QY 601 VIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKINVLKQLGTQKPPANAPSDLTIG 660
Db |||||
QY 601 VIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKINVLKQLGTQKPPANAPSDLTIG 660
Db |||||
QY 661 NEMPKYGGQSWKLAWDNPNTANGPYTLKATWTKGTGNPGERVASLVPNSLWGSILDIR 720
Db |||||
QY 661 NEMPKYGGQSWKLAWDNPNTANGPYTLKATWTKGTGNPGERVASLVPNSLWGSILDIR 720
Db |||||
QY 721 SAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDALGQGYRISGGYSLGANSYFGSSMFG 780
Db |||||
QY 721 SAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDALGQGYRISGGYSLGANSYFGSSMFG 780
Db |||||
QY 781 LAFTVEFGRSKDYVVCRSNHHACISGVYLSLTOQALCGSYLFGDAFIRASYFGNQHMKTS 840
Db |||||
QY 781 LAFTVEFGRSKDYVVCRSNHHACISGVYLSLTOQALCGSYLFGDAFIRASYFGNQHMKTS 840
Db |||||
QY 841 YTPAESDVRWNNCLAGEIGAGLPVITPSPKLYLNELRPPVQAFSVDHESFTEEGDQ 900
Db |||||
QY 841 YTPAESDVRWNNCLAGEIGAGLPVITPSPKLYLNELRPPVQAFSVDHESFTEEGDQ 900
Db |||||
QY 901 ARAFKSGHLLNLSPVGVKFDRCSTHNPKNYSFMAAYICDAVRTISGTTLLSHQETWT 960
Db |||||
QY 901 ARAFKSGHLLNLSPVGVKFDRCSTHNPKNYSFMAAYICDAVRTISGTTLLSHQETWT 960
Db |||||
QY 961 TDAFHARHGTVVVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012
Db |||||
QY 961 TDAFHARHGTVVVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012
Db |||||
```

## RESULT 2

```
US-09-542-520-2
; Sequence 2, Application US/09542520
; Patent No. 6887843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-542-520-2
```

Query Match 100.0%; Score 5267; DB 2; Length 1012;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MOTSFHKFFLSMILAYSCSLNGGYYAAEIMVPQGIYDGETLTVSFPYTVIGDPSGTVF 60
Db 1 MOTSFHKFFLSMILAYSCSLNGGYYAAEIMVPQGIYDGETLTVSFPYTVIGDPSGTVF 60
```



; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; Patent No. 6642023  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/09/612,402B  
; CURRENT FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 43  
; LENGTH: 984  
; TYPE: PR1  
; ORGANISM: Chlamydia sp.  
US-09-612-402B-43

Query Match 97.1%; Score 5116; DB 2; Length 984;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 EIMVPGIYDGETLTIVSPYTVIGDPSGTTVFSAGELTLKNDLSIAALPLSCFGLNLS 88  
DB 1 EIMVPGIYDGETLTIVSPYTVIGDPSGTTVFSAGELTLKNDLSIAALPLSCFGLNLS 60  
QY 89 FTVLGRGHSITPENIRITSTNGAALSNSAADGLFTIEGFKELSPSCNLSIAALPLPAATTNK 148  
DB 61 FTVLGRGHSITPENIRITSTNGAALSNSAADGLFTIEGFKELSPSCNLSIAALPLPAATTNK 120  
QY 149 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLNLSVSGDGAIDAKSLTVQGISKLCV 208  
DB 121 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLNLSVSGDGAIDAKSLTVQGISKLCV 180  
QY 209 QENTADGACQVTVTSFSAMANEAPIAFVANVAGVGGGIAAVQDQGGVSSSTEDP 268  
DB 181 QENTADGACQVTVTSFSAMANEAPIAFVANVAGVGGGIAAVQDQGGVSSSTEDP 240  
QY 269 VVSPSRNTAVEPDGNVARVGGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQAS 328  
DB 241 VVSPSRNTAVEPDGNVARVGGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQAS 300  
QY 329 NTSNNYGDGGAIFCKNGAAGSNNSSVSPDGGVVPFSSNVAAGKGAIYAKKLSVANC 388  
DB 301 NTSNNYGDGGAIFCKNGAAGSNNSSVSPDGGVVPFSSNVAAGKGAIYAKKLSVANC 360  
QY 389 GPVQFLRNANDGGAIVLGSSELSSADYGDIIFDGNLKRKTAKENADVNGTVSSQAI 448  
DB 361 GPVQFLRNANDGGAIVLGSSELSSADYGDIIFDGNLKRKTAKENADVNGTVSSQAI 420  
QY 449 SMGSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTDIVFANGSS 508  
DB 421 SMGSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTDIVFANGSS 480  
QY 509 TLQNVTVIEGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANLIT 568  
DB 481 TLQNVTVIEGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANLIT 540  
QY 569 LSNLHLSLSLANNVTPNTPPAQDSDHPAVTGSTAGSVTISGPIFFEDLDDTAYDR 628  
DB 541 LSNLHLSLSLANNVTPNTPPAQDSDHPAVTGSTAGSVTISGPIFFEDLDDTAYDR 600  
QY 629 YDWLGSNOKINVLKIQGTGPPANAPSDLTILGNEMPKYGGQSKLAWDPNTANNPYTL 688  
DB 601 YDWLGSNOKINVLKIQGTGPPANAPSDLTILGNEMPKYGGQSKLAWDPNTANNPYTL 660  
QY 689 KATWTKYGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDCRSYCRGLWVSGVNF 748  
DB 661 KATWTKYGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDCRSYCRGLWVSGVNF 720  
QY 749 YHRRDALGQGYRYISGGYSLGANSYFGSSMFGLAFTVEVFGRSKDYVVCRRSNHACISVY 808  
DB 721 YHRRDALGQGYRYISGGYSLGANSYFGSSMFGLAFTVEVFGRSKDYVVCRRSNHACISVY 780  
QY 809 LSTQOALCGSYLFGDAFIRASYGFGNQHMKTSTYTFABESDVRWNNCLAGIAGLPIVI 869

DB 781 LSTQOALCGSYLFGDAFIRASYGFGNQHMKTSTYTFABESDVRWNNCLAGIAGLPIVI 840  
QY 869 TPKSLYLNLRLRPVQAEFSYADHESFTTEGQARAFKSGHLNLNLVPPGVKFDRCSSSTHP 928  
DB 841 TPKSLYLNLRLRPVQAEFSYADHESFTTEGQARAFKSGHLNLNLVPPGVKFDRCSSSTHP 900  
QY 929 NKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGCVVVRGSMYASLTNSNIEVY 988  
DB 901 NKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGCVVVRGSMYASLTNSNIEVY 960  
QY 989 GHGRYEYRDASRGYGLSAGSRVRF 1012  
DB 961 GHGRYEYRDASRGYGLSAGSRVRF 984

## RESULT 4

US-09-612-402B-15  
; Sequence 15, Application US/09612402B  
; Patent No. 6642023  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; Patent No. 6642023  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/09/612,402B  
; CURRENT FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 1013  
; TYPE: PR1  
; ORGANISM: Chlamydia sp.  
US-09-612-402B-15

Query Match 97.0%; Score 5106.5; DB 2; Length 1013;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;  
QY 1 MOTSFHKFFLSMLIAYSCCLNGGGYAAEIMVPGIYDGETLTIVSPYTVIGDPSGTTVF 60  
DB 1 MOTSFHKFFLSMLIAYSCCLNGGGYAAEIMVPGIYDGETLTIVSPYTVIGDPSGTTVF 60  
QY 61 SAGELTLKNDLSIAALPLSCFGLNLSFTVLGRGHSITPENIRITSTNGAALSNSAADGL 120  
DB 61 SAGELTLKNDLSIAALPLSCFGLNLSFTVLGRGHSITPENIRITSTNGAALSNSAADGL 120  
QY 121 FTIEGFKELSPSCNLSIAALPLPAATTNKSGTSOTPTTSTPSNGTIYSKTDLLLNNEKFSF 180  
DB 121 FTIEGFKELSPSCNLSIAALPLPAATTNKSGTSOTPTTSTPSNGTIYSKTDLLLNNEKFSF 180  
QY 181 YSNLVSVDGGAIDAKSLTVQGISKLCVQFQNTAQADGGACQVTVTSFSAMANEAPIAFVAN 240  
DB 181 YSNLVSVDGGAIDAKSLTVQGISKLCVQFQNTAQADGGACQVTVTSFSAMANEAPIAFVAN 240  
QY 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVEPDGNVARVGGGIYSYGNVAF 300  
DB 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVEPDGNVARVGGGIYSYGNVAF 300  
QY 301 NNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGYDGGAIKCKNGAQ-AGSNNSSGVSFD 359  
DB 301 NNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGYDGGAIKCKNGAQ-AGSNNSSGVSFD 360  
QY 360 GEGVVPFSSNVAAGKGAIYAKKLSVANCGPVQFLRNANDGGAIVLGSSELSSADY 419  
DB 361 GEGVVPFSSNVAAGKGAIYAKKLSVANCGPVQFLRNANDGGAIVLGSSELSSADY 420  
QY 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 479  
DB 421 DMIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 480

480	Qy	NQPAQSSKLLKINDGEGYTGDIIVFANGSGSTLYQNVTTIQGRIVLREKAKLSVNSLSQTGG	539
		:	
481	Db	NQPAQSSBPLKINDGEGYTGDIIVFANGSNSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGG	540
		:	
540	Qy	SLYWEAGSTWDFVTPQPQPPAANQLITLSNLHLSSLSLLANNAVTPPTNPPQAQDSHP	599
		:	
541	Db	SLYWEAGSTLDFVTPQPQPPAANQSTLSNLHLSSLSLLANNAVTPPTNPPQAQDSHP	600
		:	
600	Qy	AVIGSTTAGSVTISGPIFEFDLDDTAYDRYDWLWGSNQKINVULKQLGTGKPPANAPSDLT	659
		:	
601	Db	AVIGSTTAGSVTISGPIFEFDLDDTAYDRYDWLWGSNQKIDVULKQLGTQPPANAPSDLT	660
		:	
660	Qy	GNEMPKGYQGGSWKLAWDPNTANNPGPYTLKATWTKTGYNPGPERVASLVPSNLWGSILDI	719
		:	
661	Db	GNEMPKGYQGGSWKLAWDPNTANNPGPYTLKATWTKTGYNPGPERVASLVPSNLWGSILDI	720
		:	
720	Qy	RSAHSAIQASVDBGRSYCRGLMWGVSVNPFYHDDRDLAQGGYRYISGGYSLGANSYFGSSMF	779
		:	
721	Db	RSAHSAIQASVDBGRSYCRGLMWGVSVNPFYHDDRDLAQGGYRYISGGYSLGANSYFGSSMF	780
		:	
780	Qy	GLAETEVFGRSKDYVVCBSNHHACIGSVYLSSTQOALCGSYLFGDAFIRASYFGNQHMKT	839
		:	
781	Db	GLAETEVFGRSKDYVVCBSNHHACIGSVYLSSTKQALCGSYVFGDAFIRASYFGNQHMKT	840
		:	
840	Qy	SYTFAESDVRWDMNCLAGEICAGLPIVITPSKLYLNELRPPVQAEFSVADHESFTEEGD	899
		:	
841	Db	SYTFAESDVCVDMNCLVGBLGVLPIVITPSKLYLNELRPPVQAEFSVADHESFTEEGD	900
		:	
900	Qy	QARAFKSGHLLNLSPVPGVKFCDRCSSTHPNKYSFMAAYICDAYRTISGTETTLTLLSHQETW	959
		:	
901	Db	QARAFRSGHLLNLSPVPGVKFCDRCSSTHPNKYSFMAAYICDAYRTISGTQITLLTLLSHQETW	960
		:	
960	Qy	TTDAFHLARHGVVVRGSGMYASLTSNIEVYHGHRVEYRDASRGYGLSAGSRVRF	1012
		:	
961	Db	TTDAFHLARHGVVVRGSGMYASLTSNIEVYHGHRVEYRDTSRGYGLSAGSKVRF	1013
		:	

## RESULT 5

```

US-09-612-402B-16
/ Sequence 16, Application US/09612402B
/ Patent No. 6642023
/ GENERAL INFORMATION:
/ APPLICANT: Jackson, W. James
/ APPLICANT: Pace, John
/ TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
/ Patent No. 6642023
/ FILE REFERENCE: 7969-086-999
/ CURRENT APPLICATION NUMBER: US/09/612,402B
/ CURRENT FILING DATE: 2000-07-06
/ PRIOR APPLICATION NUMBER: 08/942,596
/ PRIOR FILING DATE: 1997-10-02
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 16
/ LENGTH: 1013
/ TYPE: PRT
/ ORGANISM: Chlamydia sp.
US-09-612-402B-16

```

Query Match	97.0%	Score 5106.5;	DB 2;	Length 1013;
Best Local Similarity	96.7%;	Pred. No. 0;		
Matches	980;	Conservative	15; Mismatches	17; Indels
				1; Gaps
				1
Qy	1	MQTSHKFFLSMILAYSCCSLNGGYAAEIMVPGIYDGETLTVSFPYTIVGDPSGTTVF	60	
Dd	1	MQTSHKFFLSMILAYSCCSLNGGYAAEIMVPGIYDGETLTVSFPYTIVGDPSGTTVF	60	
Qy	61	SAGELTLKNLDNSTAALPLSCFGNLLGSFTVLGRGHSLTFENIRSTINGAALSAADEL	120	
Dd	61	SAGELTLKNLDNSTAALPLSCFGNLLGSFTVLGRGHSLTFENIRSTINGAALSAADEL	120	
Qy	121	FTIEGFKELFSPNCNSILLAVPAATTNKGSTPTTTPSPNGTYTSKYKTDLILLANNEKPSF	180	

121	FTIEGFKELSFNCNSLAVLPAATTNNGSQPTTTSTPNSNGTIYSKTDLLLLNNEKPSF	18
181	YSNLVSGDGAIDAISLTVOGISIKLCVFOENTAOADGACQCVWTSFSAMANEAPIAFVAN	240
181	YSNLVSGDGGTIDAISLTVOGISIKLCVFOENTAOADGACQCVWTSFSAMANEAPIAFAN	240
241	VAGVRGGGIAAVDQGGQGVSSSTSTEDPVVSFSRNTAVFPGNVARVGGIYSYGNVAPL	300
241	VAGVRGGGIAAVDQGGQGVSSSTSTEDPVVSFSRNTAVFPGNVARVGGIYSYGNVAPL	300
301	NNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNYGDGGAIFCKKCAQ--AGSNNSGSVSD	359
301	NNGKTLFLNNVASPVYIAAQPTINGQASNTSDNTGDGGAIFCKNGAQAGSNNSGSVSD	360
360	GEGVVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLRNIAIDGGAIYILGESBELSLSADYG	419
361	GEGVVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLGNIANDGGAIYILGESBELSLSADYG	420
420	DIIFDGNLKETAKENAADVNGVTVSSQAI SMGSGGKITTLRAKAGHQLFNDPIEMANGN	479
421	DIIFDGNLKETAKENAADVNGVTVSSQAI SMGSGGKITTLRAKAGHQLFNDPIEMANGN	480
480	NQPAQSSEPLKINDGEGYTGDIIVFANGNSTLYQNVTIEQRIVLREKAKLSVNSLSQTGG	539
481	NQPAQSSEPLKINDGEGYTGDIIVFANGNSTLYQNVTIEQRIVLREKAKLSVNSLSQTGG	540
540	SLYMEAGSTWDFVTVPQPPQPPAANQILITSLNHLSSLLANNNAVTPNTPPAQDSHP	599
541	SLYMEAGSTLDFVTVPQPPQPPAANQILITSLNHLSSLLANNNAVTPNTPPAQDSHP	600
600	AVLGSTTTAGSVTISGPIFFEDLDGTA VDRYDMLGSNOKI NVLKLQLGTKPPANAPSDLTL	659
601	AVLGSTTTAGSVTISGPIFFEDLDGTA VDRYDMLGSNOKI NVLKLQLGTKPPANAPSDLTL	660
660	GNEMPKYGYQGSWKLA WDPNTANNPGPYTLKATWTKGTGNPGPERVASLVPNSLWGSILDI	719
661	GNEMPKYGYQGSWKLA WDPNTANNPGPYTLKATWTKGTGNPGPERVASLVPNSLWGSILDI	720
720	RSASHAIQASVDGRSVCRLWVGVSNNPFYHDDRDLGCGYRYISGGYISLGANSYFGSSMF	779
721	RSASHAIQASVDGRSVCRLWVGVSNNPFYHDDRDLGCGYRYISGGYISLGANSYFGSSMF	780
780	GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLTOQALCGSVLFGDAFIRASYGFGNQHKMT	839
781	GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLTKQALCGSVLFGDAFIRASYGFGNQHKMT	840
840	SYTFPAGESDVRDNNCLAGHIGAGLPIVITPSKLYLNELRPFVQAEBSYADHESFTEEGD	899
841	SYTFPAGESDVRDNNCLVGBIGVGLPIVITPSKLYLNELRPFVQAEBSYADHESFTEEGD	900
900	QARAFKSGHLLNLSPVGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETTLTSHQBTW	959
901	QARAFPSGHLNLSVPVGVKFDRCSSTHPNKYSFPGAYICDAYRTISGTQTTLSHQBTW	960
960	TTDAFHLARHGTVVRGSMYASLTSNIEVYGHGRYEYRDA GRGYGLSAGSRVRF	1012
961	TTDAFHLARHGTVVRGSMYASLTSNIEVYGHGRYEYRDTSRGYGLSAGSKVRF	1013

## RESULT 6

RESULT 6  
 US-09-542-520-15  
 ; Sequence 15, Application US/09542520  
 ; Patent No. 6887843  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; APPLICANT: Pace, John L.  
 ; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF  
 ; FILE REFERENCE: 7969-076-999  
 ; CURRENT APPLICATION NUMBER: US/09/542,520  
 ; CURRENT FILING DATE: 2000-04-03  
 ; PRIOR APPLICATION NUMBER: PCT/US98/20737  
 ; PRIOR FILING DATE: 1998-10-01  
 ; NUMBER OF SEQ ID NOS: 41



601	AVIGSTTAGPVTTISGPFPEFDLDDTAYDRYDWLGSNQKIDVLKQLGTQPSANAPSDLTL	660
660	GNEMPKTYQGSWKLANDPNTANNPGPYTLKATWTKTGYNPGPERSVASLPNSLWGSILDI	719
661	GNEMPKTYQGSWKLANDPNTANNPGPYTLKATWTKTGYNPGPERSVASLPNSLWGSILDI	720
720	RSASHSATQASVDGRSYCRGLMWGVSBNFPYHDDRDLGQGYRIISGYSILGANSYFGSSMF	779
721	RSASHSATQASVDGRSYCRGLMWGVSBNFPYHDDRDLGQGYRIISGYSILGANSYFGSSMF	780
780	GLATTEVFGRSKDYVVCRSNHHACIGSVYLTSTQALCGSYLFGDAPIRASVFGNQHKMT	839
781	GLATTEVFGRSKDYVVCRSNHHACIGSVYLTSTQALCGSYLFGDAPIRASVFGNQHKMT	840
840	SYTFAESDVRWNNKLAGEICAGLPIVITPSSKLYLNEIRFPVQAEFYSADHESFTEGSD	899
841	SYTFAESDVRWNNKLAGEICAGLPIVITPSSKLYLNEIRFPVQAEFYSADHESFTEGSD	900
900	QARAFKGGHLLNLSPVGVGKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	959
901	QARAFRSGHLNLSVPVGVGKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	960
960	TTDAFLARHGVIYVRGSGMYASLTNSIEVYGHGRYEYRDASRGYGLSAGSRVRP	1012
961	TTDAFLARHGVIYVRGSGMYASLTNSIEVYGHGRYEYRTSGYGLSAGSKVRP	1013

## RESULT 8

```

US-09-556-877-190
/ Sequence 190, Application US/09556877
/ Patent No. 6432916
/ GENERAL INFORMATION:
/ APPLICANT: Probst, Peter
/ APPLICANT: Bhatia, Ajay
/ APPLICANT: Skeiky, Yasir
/ APPLICANT: Fling, Steve
/ APPLICANT: Malsonneuve, Jeff
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
/ TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
/ FILE REFERENCE: 210121.469C5
/ CURRENT APPLICATION NUMBER: US/09/556,877
/ CURRENT FILING DATE: 2000-04-19
/ NUMBER OF SEQ ID NOS: 305
/ SOFTWARE: FastSEQ for Windows Version 3.0/4.0
/ SEQ ID NO 190
/ LENGTH: 1006
/ TYPE: PRT
/ ORGANISM: Chlamydia
US-09-556-877-190

```

	Query Match	96.6%	Score 5090;	DB 2;	Length 1006;
	Best Local Similarity	99.7%;	pred. No. 0;		
	Matches 979;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	31	MVPGQIYDGETLTVSPPYTWIGDSPGCTTVPSAGELTLKLNLDNSTAALPLSCFGNLLGSPT	90		
Db	25	MI PGQIYDGETLTVSPPYTWIGDSPGCTTVPSAGELTLKLNLDNSTAALPLSCFGNLLGSPT	84		
Qy	91	VLGRGHSLTPIENIRTSNTGAAALNSAADGLFTIEGPKELSPNCNSLLAVLPAATTNKG	150		
Db	85	VLGRGHSLTPIENIRTSNTGAAALNSAADGLFTIEGPKELSPNCNSLLAVLPAATTNKG	144		
Qy	151	QTPTTSTPNGTITSKTDLLLLNNKXPSFYNLVSGDGGDAIDAKSLTVQGISKLCVQOE	210		
Db	145	QTPTTSTPNGTITSKTDLLLLNNKXPSFYNLVSGDGGDAIDAKSLTVQGISKLCVQOE	204		
Qy	211	NTAQADGACQVVTSPSAMANEAPIAVANVAVYRGGGIAVQDQGGVSSSSTSTEDPVV	270		
Db	205	NTAQADGACQVVTSPSAMANEAPIAVANVAVYRGGGIAVQDQGGVSSSSTSTEDPVV	264		
Qy	271	SPSRNTAVEFDGNVARYVGGGIYSYNVAFLNNGKTLFLNNVASFVYIAAKQPTSGQASNT	330		

Db	265	SFSRNTAVEFDGNVARVGGGIYSYGNVAFNLNKGKTLFLNNNVASPVYIAAKQPTSQASNT	324
Qy	331	SNNTYGDGGAIFCKNGAQAGSNNSGSVFDGEGVVPFSSNVNVAAGKGGAIYAKKLSVANCGP	390
Db	325	SNNTYGDGGAIFCKNGAQAGSNNSGSVFDGEGVVPFSSNVNVAAGKGGAIYAKKLSVANCGP	384
Qy	391	VQFLRNTANDGGAITYLGBESGBLSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM	450
Db	385	VQFLRNTANDGGAITYLGBESGBLSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM	444
Qy	451	GSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL	510
Db	445	GSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL	504
Qy	511	YQNTYIEQGRIVLREKAKLSVNSLSQTGSSLYMEAGSTWDFVTPOPPPOPPAANQILITLS	570
Db	505	YQNTYIEQGRIVLREKAKLSVNSLSQTGSSLYMEAGSTWDFVTPOPPPOPPAANQILITLS	564
Qy	571	NLHLSLSLLANNAVTPPTNPPAQDSHPAIVGSTTAGSVTISGPIFFPEDLDDTAYDRYD	630
Db	565	NLHLSLSLLANNAVTPPTNPPAQDSHPAIVGSTTAGSVTISGPIFFPEDLDDTAYDRYD	624
Qy	631	WLGSNQKINVILQLGTQKPPANAPSDTLTGNEMPKYGYQGSWKLAWDNTANNNGPYTLKA	690
Db	625	WLGSNQKINVILQLGTQKPPANAPSDTLTGNEMPKYGYQGSWKLAWDNTANNNGPYTLKA	684
Qy	691	TWTKTGNPGERVASILVPNSLWGSIIIDIRSAHSAIQASVDGRSVCRLWVSGVSNPFYH	750
Db	685	TWTKTGNPGERVASILVPNSLWGSIIIDIRSAHSAIQASVDGRSVCRLWVSGVSNPFYH	744
Qy	751	DRDALGQGYRYISGGYSLGANSYFGSSNMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS	810
Db	745	DRDALGQGYRYISGGYSLGANSYFGSSNMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS	804
Qy	811	TQOALCGSYLFGDAPIRASYPFGNQHMKTISTYFAEBSDDVRDNNCLAGEIGAGLPVITP	870
Db	805	TQOALCGSYLFGDAPIRASYPFGNQHMKTISTYFAEBSDDVRDNNCLAGEIGAGLPVITP	864
Qy	871	SKLYLNELRPPVQAFSPYADHESFTEEGDQARAFKSGHLLNLSVPVGVKFPDRCSTHPNK	930
Db	865	SKLYLNELRPPVQAFSPYADHESFTEEGDQARAFKSGHLLNLSVPVGVKFPDRCSTHPNK	924
Qy	931	YSFMAAYICDAYRTISGTETTTLLSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	990
Db	925	YSFMAAYICDAYRTISGTETTTLLSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	984
Qy	991	GRYBYRDASRGYGLSAGSRVRP	1012
Db	985	GRYBYRDASRGYGLSAGSKVRP	1006

```

RESULT 9
US-09-620-412C-190
; Sequence 190, Application US/09620412C
; Patent NO. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

```

Query Match 96.6%; Score 5090; DB 2; Length 1006;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 979; Conservative 2; Mismatches 1; Indels 0; Gaps 0

```
QY 31 MYPOGIYDGETLVSPFYTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFNLIGSFT 90
DB 1:|||||YDGETLVSPFYTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFNLIGSFT 84
QY 91 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 150
DB 85 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 144
QY 151 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVQ 210
DB 145 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVQ 204
QY 211 NTAQADGGACQVVTFSAMANEAPAFVANVAVRGGGIAAVQDGGQVSSSTSTEDPV 270
DB 205 NTAQADGGACQVVTFSAMANEAPAFVANVAVRGGGIAAVQDGGQVSSSTSTEDPV 264
QY 271 SFSRNTAVFDCGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQ 330
DB 265 SFSRNTAVFDCGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQ 324
QY 331 SNNGDGGAI FCKNGAQAGNNSSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVAN 390
DB 325 SNNGDGGAI FCKNGAQAGNNSSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVAN 384
QY 391 VOFLRNANDGGAIIYLGESGELSLSDYGDIIIPDGNLKTAKENADVNGVTVSSQ 450
DB 385 VOFLRNANDGGAIIYLGESGELSLSDYGDIIIPDGNLKTAKENADVNGVTVSSQ 444
QY 451 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANG 510
DB 445 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANG 504
QY 511 YQNVTTIEQRIIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPQPPAANQ 570
DB 505 YQNVTTIEQRIIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPQPPAANQ 564
QY 571 NLHLSLSSLLANNAVTPNPPAQDHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDR 630
DB 565 NLHLSLSSLLANNAVTPNPPAQDHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDR 624
QY 631 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNP 690
DB 625 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNP 684
QY 691 TWTKGTGYNPQPERVASILVPSNLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGV 750
DB 685 TWTKGTGYNPQPERVASILVPSNLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGV 744
QY 751 YSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGTVVGRGMYASLTNIEV 930
DB 925 YSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGTVVGRGMYASLTNIEV 984
QY 991 GRYEYRDASRGYGLSAGSRVP 1012
DB 985 GRYEYRDASRGYGLSAGSRVP 1006
```

RESULT 10

US-09-598-419-190

; Sequence 190, Application US/09598419

```
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-190
```

```
Query Match 96.6%; Score 5090; DB 2; Length 1006;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 979; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 31 MYPOGIYDGETLVSPFYTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFNLIGSFT 90
DB 25 MIPOGIYDGETLVSPFYTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFNLIGSFT 84
QY 91 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 150
DB 85 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 144
QY 151 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVQ 210
DB 145 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVQ 204
QY 211 NTAQADGGACQVVTFSAMANEAPAFVANVAVRGGGIAAVQDGGQVSSSTSTEDPV 270
DB 205 NTAQADGGACQVVTFSAMANEAPAFVANVAVRGGGIAAVQDGGQVSSSTSTEDPV 264
QY 271 SFSRNTAVFDCGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQ 330
DB 265 SFSRNTAVFDCGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQ 324
QY 331 SNNGDGGAI FCKNGAQAGNNSSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVAN 390
DB 325 SNNGDGGAI FCKNGAQAGNNSSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVAN 384
QY 391 VOFLRNANDGGAIIYLGESGELSLSDYGDIIIPDGNLKTAKENADVNGVTVSSQ 450
DB 385 VOFLRNANDGGAIIYLGESGELSLSDYGDIIIPDGNLKTAKENADVNGVTVSSQ 444
QY 451 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANG 510
DB 445 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANG 504
QY 511 YQNVTTIEQRIIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPQPPAANQ 570
DB 505 YQNVTTIEQRIIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPQPPAANQ 564
QY 571 NLHLSLSSLLANNAVTPNPPAQDHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDR 630
DB 565 NLHLSLSSLLANNAVTPNPPAQDHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDR 624
QY 631 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNP 690
DB 625 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNP 684
QY 691 TWTKGTGYNPQPERVASILVPSNLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGV 750
DB 685 TWTKGTGYNPQPERVASILVPSNLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGV 744
QY 751 DRDALGGYRYISGGYSLGANSYFGSSMFLAFTFVFGRSKDYVVCRSNHHACIGSV 810
DB 745 DRDALGGYRYISGGYSLGANSYFGSSMFLAFTFVFGRSKDYVVCRSNHHACIGSV 804
```

QY 811 TQALCGSYLFGDAFIRASYGFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 870  
DB 805 TQALCGSYLFGDAFIRASYGFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 864  
QY 871 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLLNLSPVGVKFDRCSTTHPNK 930  
DB 865 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLLNLSPVGVKFDRCSTTHPNK 924  
QY 931 YSFMAAYICDAYRTISGTETLLSHOETWTTDAFHARHGVVVRGSMYASLTSNIEVYGH 990  
DB 925 YSFMAAYICDAYRTISGTETLLSHOETWTTDAFHARHGVVVRGSMYASLTSNIEVYGH 984  
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012  
DB 985 GRYEYRDASRGYGLSAGSKVRF 1006  
RESULT 11  
US-09-556-877-176  
; Sequence 176, Application US/09556877  
; Patent No. 6432916  
; GENERAL INFORMATION:  
; APPLICANT: Probst, Peter  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Fling, Steve  
; APPLICANT: Maisonneuve, Jeff  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C5  
; CURRENT APPLICATION NUMBER: US/09/556,877  
; CURRENT FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 305  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 176  
; LENGTH: 982  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; FEATURES:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(982)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-556-877-176

Query Match 96.5%; Score 5084; DB 2; Length 982;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 31 MVPOGIYDGETLTVSPFVTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 90  
DB 1 MIPOGIYDGETLTVSPFVTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 60  
QY 91 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGKELSFSCNCSLLAVLPAATTKGS 150  
DB 61 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGKELSFSCNCSLLAVLPAATTKGS 120  
QY 151 QTPTTTSPNGTIYSKNTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVPOE 210  
DB 121 QTPTTTSPNGTIYSKNTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVPOE 180  
QY 211 NTAQDGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVODQOQGVSSSTEDPVV 270  
DB 181 NTAQDGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVODQOQGVSSSTEDPVV 240  
QY 271 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 330  
DB 241 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 300  
QY 331 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFPFSSNVAAGKGAIAKYLKLVANCGP 390  
DB 301 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFPFSSNVAAGKGAIAKYLKLVANCGP 360  
QY 391 VOPLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENADVNGVTVSSQAISM 450

DB 361 VOPLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENADVNGVTVSSQAISM 420  
QY 451 GSGGKIITTLRAKAGHQLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 510  
DB 421 GSGGKIITTLRAKAGHQLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 480  
QY 511 YQNTYISQGRIVLREKAKLSVNSLSQSGSYMEAGSTMDPVTPQPPQPPAANQLITLS 570  
DB 481 YQNTYISQGRIVLREKAKLSVNSLSQSGSYMEAGSTMDPVTPQPPQPPAANQLITLS 540  
QY 571 NLHLSLSLLANNAVTPPTPPAQDSHPAIVIGSTTAGSVTISGPIFFEDLDDTAYDRYD 630  
DB 541 NLHLSLSLLANNAVTPPTPPAQDSHPAIVIGSTTAGSVTISGPIFFEDLDDTAYDRYD 600  
QY 631 WLGSNQKINVLKQLGKTPPANAPSDTLTGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 690  
DB 601 WLGSNQKINVLKQLGKTPPANAPSDTLTGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 660  
QY 691 TWTGTGYNPGPERVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNRPYH 750  
DB 661 TWTGTGYNPGPERVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNRPYH 720  
QY 751 DRDALGQGYRYISGYSILGANSYFGSSMFGLAFTTEVFGRSKDYVVCRSNHHACIGSVYLS 810  
DB 721 DRDALGQGYRYISGYSILGANSYFGSSMFGLAFTTEVFGRSKDYVVCRSNHHACIGSVYLS 780  
QY 811 TQALCGSYLFGDAFIRASYGFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 870  
DB 781 TQALCGSYLFGDAFIRASYGFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 840  
QY 871 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLLNLSPVGVKFDRCSTTHPNK 930  
DB 841 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLLNLSPVGVKFDRCSTTHPNK 900  
QY 931 YSFMAAYICDAYRTISGTETLLSHOETWTTDAFHARHGVVVRGSMYASLTSNIEVYGH 990  
DB 901 YSFMAAYICDAYRTISGTETLLSHOETWTTDAFHARHGVVVRGSMYASLTSNIEVYGH 960  
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012  
DB 961 GRYEYRDASRGYGLSAGSKVRF 982  
RESULT 12  
US-09-620-412C-176  
; Sequence 176, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 176  
; LENGTH: 982  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; FEATURES:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(982)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-620-412C-176  
Query Match 96.5%; Score 5084; DB 2; Length 982;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 31 MVPOGIYDGETLTVSPFVTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 90  
DB 1 MIPOGIYDGETLTVSPFVTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 60  
QY 91 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGKELSFSCNCSLLAVLPAATTKGS 150  
DB 61 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGKELSFSCNCSLLAVLPAATTKGS 120  
QY 151 QTPTTTSPNGTIYSKNTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVPOE 210  
DB 121 QTPTTTSPNGTIYSKNTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVPOE 180  
QY 211 NTAQDGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVODQOQGVSSSTEDPVV 270  
DB 181 NTAQDGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVODQOQGVSSSTEDPVV 240  
QY 271 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 330  
DB 241 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 300  
QY 331 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFPFSSNVAAGKGAIAKYLKLVANCGP 390  
DB 301 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFPFSSNVAAGKGAIAKYLKLVANCGP 360  
QY 391 VOPLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENADVNGVTVSSQAISM 450

```
Db 1 MIPQGIYDGETLTVSPFYTVIGDPGSGTTVFSAGELTLKNLDSIAALPLSCFNLGSGFT 60
Qy 91 VLGRGHSUTPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 150
Db 61 VLGRGHSUTPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 120
Qy 151 QTPPTTSPNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTQVQISKLQVQFE 210
Db 121 QTPPTTSPNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTQVQISKLQVQFE 180
Qy 211 NTAQADGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 270
Db 181 NTAQADGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 240
Qy 271 SPSRNTAVEFDGNVARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGASNT 330
Db 241 SPSRNTAVEFDGNVARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGASNT 300
Qy 331 SNNGYDGGAI FCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGAIVAKKLSVANCGP 390
Db 301 SNNGYDGGAI FCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGAIVAKKLSVANCGP 360
Qy 391 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 450
Db 361 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 420
Qy 451 GSGGKIITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
Db 421 GSGGKIITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 480
Qy 511 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 570
Db 481 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 540
Qy 571 NLHLSLSSLLANNAVTPNPPAODSHPAVIGSTTAGSVTTISGPIFFEDLDLDTAYDRYD 630
Db 541 NLHLSLSSLLANNAVTPNPPAODSHPAVIGSTTAGSVTTISGPIFFEDLDLDTAYDRYD 600
Qy 631 WLGSNQKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 690
Db 601 WLGSNQKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 660
Qy 691 TWTKTGYNPGERVASLVPNSLWGSITLDRSAHSAIQASVDGGRSYCRGLWVSGVSNFFYH 750
Db 661 TWTKTGYNPGERVASLVPNSLWGSITLDRSAHSAIQASVDGGRSYCRGLWVSGVSNFFYH 720
Qy 751 DRDALGQGYRIYISGGYSLGANSYFGSGMFLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 810
Db 721 DRDALGQGYRIYISGGYSLGANSYFGSGMFLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 780
Qy 811 TQOALCGSYLFGDAFIRASYGFGNQHMKTSYTPAESDVVRWNNCLAGEIGAGLPVITP 870
Db 781 TQOALCGSYLFGDAFIRASYGFGNQHMKTSYTPAESDVVRWNNCLAGEIGAGLPVITP 840
Qy 871 SKLYNELRPVQAEFSAADHESFTEGQOARAFKSHLLNLSPVGVKFDRCSSHTPNK 930
Db 841 SKLYNELRPVQAEFSAADHESFTEGQOARAFKSHLLNLSPVGVKFDRCSSHTPNK 900
Qy 931 YSPMAAYICDAYRTISGTETTLTSHOBTWTTDAPHLARHGCVVRGSMYASLTNSIEVYGH 990
Db 901 YSPMAAYICDAYRTISGTETTLTSHOBTWTTDAPHLARHGCVVRGSMYASLTNSIEVYGH 960
Qy 991 GRYEYRDASRGYGLSAGSRVRF 1012
Db 961 GRYEYRDASRGYGLSAGSKVXF 982
```

## RESULT 13

US-09-598-419-176

; Sequence 176, Application US/09598419

; Patent No. 6565856

; GENERAL INFORMATION:

; APPLICANT: Skelky, Yasir A.W.

```
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-598-419-176
```

Query Match 96.5%; Score 5084; DB 2; Length 982;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy 31 MIPQGIYDGETLTVSPFYTVIGDPGSGTTVFSAGELTLKNLDSIAALPLSCFNLGSGFT 90
Db 1 MIPQGIYDGETLTVSPFYTVIGDPGSGTTVFSAGELTLKNLDSIAALPLSCFNLGSGFT 60
Qy 91 VLGRGHSUTPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 150
Db 61 VLGRGHSUTPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 120
Qy 151 QTPPTTSPNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTQVQISKLQVQFE 210
Db 121 QTPPTTSPNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTQVQISKLQVQFE 180
Qy 211 NTAQADGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 270
Db 181 NTAQADGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 240
Qy 271 SPSRNTAVEFDGNVARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGASNT 330
Db 241 SPSRNTAVEFDGNVARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGASNT 300
Qy 331 SNNGYDGGAI FCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGAIVAKKLSVANCGP 390
Db 301 SNNGYDGGAI FCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGAIVAKKLSVANCGP 360
Qy 391 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 450
Db 361 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 420
Qy 451 GSGGKIITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
Db 421 GSGGKIITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 480
Qy 511 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 570
Db 481 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 540
Qy 571 NLHLSLSSLLANNAVTPNPPAODSHPAVIGSTTAGSVTTISGPIFFEDLDLDTAYDRYD 630
Db 541 NLHLSLSSLLANNAVTPNPPAODSHPAVIGSTTAGSVTTISGPIFFEDLDLDTAYDRYD 600
Qy 631 WLGSNQKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 690
Db 601 WLGSNQKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 660
Qy 691 TWTKTGYNPGERVASLVPNSLWGSITLDRSAHSAIQASVDGGRSYCRGLWVSGVSNFFYH 750
Db 661 TWTKTGYNPGERVASLVPNSLWGSITLDRSAHSAIQASVDGGRSYCRGLWVSGVSNFFYH 720
Qy 751 DRDALGQGYRIYISGGYSLGANSYFGSGMFLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 810
Db 721 DRDALGQGYRIYISGGYSLGANSYFGSGMFLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 780
```



QY 811 TQOALCGSYLFGDAPIRASYGFGNQHKTSYTFABESDVRDNNCLAGEIGAGLPIVITP 870  
Db 781 TQOALCGSYLFGDAPIRASYGFGNQHKTSYTFABESDVRDNNCLAGEIGAGLPIVITP 840  
QY 871 SKLYLNELRPPVQAEFSYADHESFTEBGDQARAFKSGHLLNLSPVGVKFDRCSTHFNK 930  
Db 841 SKLYLNELRPPVQAEFSYADHESFTEBGDQARAFKSGHLLNLSPVGVKFDRCSTHFNK 900  
QY 931 YSFMAYICDAVRTISGTETTLSSHQETWTTDAPFLARHGVVVRGSMYASLTNSIEVYGH 990  
Db 901 YSFMAYICDAVRTISGTETTLSSHQETWTTDAPFLARHGVVVRGSMYASLTNSIEVYGH 960  
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012  
Db 961 GRYEYRDASRGYGLSAGSKVXF 982

## RESULT 14

US-10-197-220-169

; Sequence 169, Application US/10197220

; Patent No. 6919187

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Guderian, Jeff

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Maisonneuve, Jean-Francois L.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

; FILE REFERENCE: 210121.515C3

; CURRENT APPLICATION NUMBER: US/10/197,220

; CURRENT FILING DATE: 2002-07-15

; NUMBER OF SEQ ID NOS: 175

; SEQ ID NO 169

; LENGTH: 670

; TYPE: PRT

; ORGANISM: Chlamydia

US-10-197-220-169

Query Match 63.3%; Score 3336.5; DB 2; Length 670;  
Best Local Similarity 97.2%; Pred. No. 3.1e-283;  
Matches 651; Conservative 11; Mismatches 7; Indels 1; Gaps 1;  
QY 28 AEIMVPOGIYDGETLTVSFPYTVIGDPGSGTTFVSAGELTLKNLNSIAALPLSCFNGLLG 87  
Db 1 AEIMVPOGIYDGETLTVSFPYTVIGDPGSGTTFVSAGELTLKNLNSIAALPLSCFNGLLG 60  
QY 88 SFTVLGRGHSITFENIRSTNGAALNSAAGDLFTIEGFKELSPFNCNLSLLAVLPAATTN 147  
Db 61 SFTVLGRGHSITFENIRSTNGAALNSAAGDLFTIEGFKELSPFNCNLSLLAVLPAATTN 120  
QY 148 KGSQPTTTTSPNGTIYSKTDLLLNNEKFSFYNLSVGDGGAIDAKSLTVQGISKLCV 207  
Db 121 KGSQPTTTTSPNGTIYSKTDLLLNNEKFSFYNLSVGDGGAIDAKSLTVQGISKLCV 180  
QY 208 FOENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGGIAAVDQCGQVSSSTSTED 267  
Db 181 FOENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGGIAAVDQCGQVSSSTSTED 240  
QY 268 PVVPSRNTAVFPGDNVARVGGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQQA 327  
Db 241 PVVPSRNTAVFPGDNVARVGGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQQA 300  
QY 328 SNTSNYGDGGAIFCKNGAQ-AGSNNGSVSFDGEGVVFSSNVAAGKGGAIYAKKLSVA 386  
Db 301 SNTSNYGDGGAIFCKNGAQAGSNNGSVSFDGEGVVFSSNVAAGKGGAIYAKKLSVA 360  
QY 387 NCGPVQFLRNTANDGGAIIYLGESGELSLSADYGDIIIFDGNLKRTAKENAAVNGVTVSSQ 446  
Db 361 NCGPVQFLRNTANDGGAIIYLGESGELSLSADYGDIIIFDGNLKRTAKENAAVNGVTVSSQ 420  
QY 447 AISMGSGGKITTIRAKAGHOILFNDPIEMANGNQPASSEPLKINDGEGYTGDIVPANG 506

Db 421 AISMGSGKITTIRAKAGHOILFNDPIEMANGNQPASSEPLKINDGEGYTGDIVPANG 480  
QY 507 SSTIYQNVITIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPPQPQPPAANQL 566  
Db 481 NSTLIYQNVITIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPPQPQPPAANQL 540  
QY 567 ITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626  
Db 541 ITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 600  
QY 627 DRYDMLGSNOKINVLKLGTKPPANAPSDLTILGNEMPKYQYQGSWKLAWDPNTANNQPY 686  
Db 601 DRYDMLGSNOKIDVLKLGTKPPANAPSDLTILGNEMPKYQYQGSWKLAWDPNTANNQPY 660  
QY 687 TLKATWTKTG 696  
Db 661 TLKATWTKTG 670

## RESULT 15

US-09-612-402B-17

; Sequence 17, Application US/09612402B

; Patent No. 6642023

; GENERAL INFORMATION:

; APPLICANT: Jackson, W. James

; APPLICANT: Pace, John

; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof

; Patent No. 6642023

; FILE REFERENCE: 7969-086-999

; CURRENT APPLICATION NUMBER: US/09/612,402B

; CURRENT FILING DATE: 2000-07-06

; PRIOR APPLICATION NUMBER: 08/942,596

; PRIOR FILING DATE: 1997-10-02

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 17

; LENGTH: 505

; TYPE: PRT

; ORGANISM: Chlamydia sp.

US-09-612-402B-17

Query Match 48.5%; Score 2552; DB 2; Length 505;  
Best Local Similarity 100.0%; Pred. No. 1.1e-214;  
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 EIMVPOGIYDGETLTVSFPYTVIGDPGSGTTFVSAGELTLKNLNSIAALPLSCFNGLLG 88  
Db 1 EIMVPOGIYDGETLTVSFPYTVIGDPGSGTTFVSAGELTLKNLNSIAALPLSCFNGLLG 60  
QY 89 FTVLGRGHSITFENIRSTNGAALNSAAGDLFTIEGFKELSPFNCNLSLLAVLPAATTN 148  
Db 61 FTVLGRGHSITFENIRSTNGAALNSAAGDLFTIEGFKELSPFNCNLSLLAVLPAATTN 120  
QY 149 GSOPTTTTSPNGTIYSKTDLLLNNEKFSFYNLSVGDGGAIDAKSLTVQGISKLCV 208  
Db 121 GSOPTTTTSPNGTIYSKTDLLLNNEKFSFYNLSVGDGGAIDAKSLTVQGISKLCV 180  
QY 209 QENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGGIAAVDQCGQVSSSTSTEDP 268  
Db 181 QENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGGIAAVDQCGQVSSSTSTEDP 240  
QY 269 VVPSRNTAVFPGDNVARVGGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQQA 328  
Db 241 VVPSRNTAVFPGDNVARVGGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQQA 300  
QY 329 NTSNNGYDGGAIIFCKNGAQAGSNNGSVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 388  
Db 301 NTSNNGYDGGAIIFCKNGAQAGSNNGSVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 360  
QY 389 GPVQFLRNTANDGGAIIYLGESGELSLSADYGDIIIFDGNLKRTAKENAAVNGVTVSSQAI 448  
Db 361 GPVQFLRNTANDGGAIIYLGESGELSLSADYGDIIIFDGNLKRTAKENAAVNGVTVSSQAI 420

QY 449 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 508  
DB 421 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 480  
QY 509 TLYQNVTIEQRIIVLRKAKLSVNS 533  
DB 481 TLYQNVTIEQRIIVLRKAKLSVNS 505

RESULT 16  
US-09-542-520-17  
; Sequence 17, Application US/09542520  
; Patent No. 6887843  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John L.  
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 7969-076-999  
; CURRENT APPLICATION NUMBER: US/09/542,520  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/US98/20737  
; PRIOR FILING DATE: 1998-10-01  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-09-542-520-17

Query Match 48.4%; Score 2547; DB 2; Length 505;  
Best Local Similarity 99.8%; Pred. No. 3.1e-214; Indels 0; Gaps 0;  
Matches 504; Conservative 1; Mismatches 0;

QY 29 EIMVPGIYDGETLTVSPFTVIGDPSGTTVFSAGELTLKNDLSIAALPLSCFGLLGS 88  
DB 1 EIMVPGIYDGETLTVSPFTVIGDPSGTTVFSAGELTLKNDLSIAALPLSCFGLLGS 60

QY 89 FTVLGRHSLTFENIRFTSTNGAALSNSADGLFTTIGFKELSPSNCNSLLAVLPAATTK 148  
DB 61 FTVLGRHSLTFENIRFTSTNGAALSNSADGLFTTIGFKELSPSNCNSLLAVLPAATTK 120

QY 149 GSQTPTTSTPSNGTIYSKTDLLANNEKESFYSLVSGDGGAIKAKSLTVQGISKLCVF 208  
DB 121 GSQTPTTSTPSNGTIYSKTDLLANNEKESFYSLVSGDGGAIKAKSLTVQGISKLCVF 180

QY 209 QENTAOADGACQVWTSFSAMANEAPAFVANVAGVRGGIAAVQDQGGVSSSTSTEDP 268  
DB 181 QENTAOADGACQVWTSFSAMANEAPAFVANVAGVRGGIAAVQDQGGVSSSTSTEDP 240

QY 269 VVSFSRNTAYEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQAS 328  
DB 241 VVSFSRNTAYEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQAS 300

QY 329 NTSNNYDGGAIKCKNGAQAGSNNSSVSFDGEGVFPSSNVAAGKGAIVAKKLSVANC 388  
DB 301 NTSNNYDGGAIKCKNGAQAGSNNSSVSFDGEGVFPSSNVAAGKGAIVAKKLSVANC 360

QY 389 GPVQFLNIANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI 448  
DB 361 GPVQFLNIANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI 420

QY 449 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 508  
DB 421 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 480

QY 509 TLYQNVTIEQRIIVLRKAKLSVNS 533  
DB 481 TLYQNVTIEQRIIVLRKAKLSVNS 505

RESULT 17  
US-09-612-402B-36

; Sequence 36, Application US/09612402B  
; Patent No. 6642023  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; Patent No. 6642023  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/09/612,402B  
; CURRENT FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 36  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-09-612-402B-36

Query Match 44.6%; Score 2350; DB 2; Length 458;  
Best Local Similarity 100.0%; Pred. No. 5e-197;  
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GGACQVWTSFSAMANEAPAFVANVAGVRGGIAAVQDQGGVSSSTSTEDPVSFSRNT 276  
DB 1 GGACQVWTSFSAMANEAPAFVANVAGVRGGIAAVQDQGGVSSSTSTEDPVSFSRNT 60

QY 277 AVEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNTSNYGD 336  
DB 61 AVEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNTSNYGD 120

QY 337 GCAIFCKNGAQAGSNNSSVSFDGEGVFPSSNVAAGKGAIVAKKLSVANGCPVQFLRN 396  
DB 121 GCAIFCKNGAQAGSNNSSVSFDGEGVFPSSNVAAGKGAIVAKKLSVANGCPVQFLRN 180

QY 397 IANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI SMGSGKI 456  
DB 181 IANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI SMGSGKI 240

QY 457 TTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSSSTLYQNVTI 516  
DB 241 TTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSSSTLYQNVTI 300

QY 517 EGRIVLRKAKLSVNSLSQTSGLYMEAGSTWDFVTPQPPQPPAANQLITLSNHLSL 576  
DB 301 EGRIVLRKAKLSVNSLSQTSGLYMEAGSTWDFVTPQPPQPPAANQLITLSNHLSL 360

QY 577 SSLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNQ 636  
DB 361 SSLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNQ 420

QY 637 KINVLKQLGTGKPPANAPSDLTIGNEMPKYQGSWKL 674  
DB 421 KINVLKQLGTGKPPANAPSDLTIGNEMPKYQGSWKL 458

RESULT 18  
US-09-542-520-36  
; Sequence 36, Application US/09542520  
; Patent No. 6887843  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John L.  
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 7969-076-999  
; CURRENT APPLICATION NUMBER: US/09/542,520  
; CURRENT FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: PCT/US98/20737  
; PRIOR FILING DATE: 1998-10-01  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36

```
/ LENGTH: 458
/ TYPE: PRT
/ ORGANISM: Chlamydia
US-09-542-520-36

Query Match      44.6%; Score 2350; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 5e-197;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GGACQVTSFSAWANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNT 276
Db 1 GGACQVTSFSAWANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNT 60
QY 277 AVEFGNVARVGGGYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 336
Db 61 AVEFGNVARVGGGYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 120
QY 337 GGAIFCKNGAQAQAGNNSSVSFDDGGVVFSSNVAAGKGAIYAKKLSVANGCPVQFLRN 396
Db 121 GGAIFCKNGAQAQAGNNSSVSFDDGGVVFSSNVAAGKGAIYAKKLSVANGCPVQFLRN 180
QY 397 TANDGGAAYLGBSGLSLADYDGIIPDNLKRKTAKENAADVGVTVSSQAISMGGGKI 456
Db 181 TANDGGAAYLGBSGLSLADYDGIIPDNLKRKTAKENAADVGVTVSSQAISMGGGKI 240
QY 457 TTRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTDIVFANGSSTLYQNTI 516
Db 241 TTRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTDIVFANGSSTLYQNTI 300
QY 517 EGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPQPPQPPAANQLITLSLHLSL 576
Db 301 EGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPQPPQPPAANQLITLSLHLSL 360
QY 577 SLLANNAVTPNTPPAQDSHPAVIGSTTAGSVTISGPIPFDDDDTAYDRYDLWLSNQ 636
Db 361 SLLANNAVTPNTPPAQDSHPAVIGSTTAGSVTISGPIPFDDDDTAYDRYDLWLSNQ 420
QY 637 KINVLKLOLGTGKPPANAPSDLTGLNEMPKYGYQGSWKL 674
Db 421 KINVLKLOLGTGKPPANAPSDLTGLNEMPKYGYQGSWKL 458

RESULT 19
US-09-612-402B-37
; Sequence 37, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT FILING DATE: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-37

Query Match      32.9%; Score 1735; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.3e-143;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 747
Db 1 LKATWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 60
QY 748 FYHDDRDLGQGYRYISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHHACIGSV 807
Db 61 FYHDDRDLGQGYRYISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHHACIGSV 120
QY 808 YLSTQOALCGSYLFGDAFIRASYGFGNQHKMTSYTFAEESDVRDNNCLAGEICAGLPV 867
Db 121 YLSTQOALCGSYLFGDAFIRASYGFGNQHKMTSYTFAEESDVRDNNCLAGEICAGLPV 180
QY 868 ITPSKLYLNELRPFVQAEPFSAHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 927
Db 181 ITPSKLYLNELRPFVQAEPFSAHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 240
QY 928 PNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAFHARHGTVVGRGMSYASLTSTNIEV 987
Db 241 PNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAFHARHGTVVGRGMSYASLTSTNIEV 300
QY 988 YGHGRYEYRDASRGYGLSAGSRVRF 1012
Db 301 YGHGRYEYRDASRGYGLSAGSRVRF 325

RESULT 21
US-09-438-185A-455
; Sequence 455, Application US/09438185A

Query Match      32.9%; Score 1735; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.3e-143;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 747
Db 1 LKATWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 60
QY 748 FYHDDRDLGQGYRYISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHHACIGSV 807
```

```
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 455
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0453
; US-09-438-185A-455

Query Match      25.9%; Score 1365.5; DB 2; Length 999;
Best Local Similarity 33.5%; Pred. No. 3.8e-110;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY      1  MOTSFHKFPLSMILAYSCCSLNGGVAABIMVQGIYDGETLTVSPYTVIGPSPGTVTF 60
Db      1  MTSIRKFLISTTLAPCFAST---AFTVEVIMSENFDDGSSGKI-PPYTTLSDFRGLTLCI 82
QY      61  SAGELTKNLDNSIAALPLSCFNLGSGFTVLGRGHSLTPENIRTSNNGAALSNAADG- 119
Db      61  FSGDLIYANLDNAISRTSSCFNRRAGALQILKGKGVFSLNIRSSADGAIAISSVITQNP 142
QY      120  ---LFTIEGKELSPFNCNSLLAVLPAATNKGSOPTTTTSTPSNGTIYKTDLLLNNE 176
Db      120  ---LFTIEGKELSPFNCNSLLAVLPAATNKGSOPTTTTSTPSNGTIYKTDLLLNNE 176
QY      143  ELCPLSFSGFSQMFONCESLT-----SDTSASNVI PHASAIYATTPMLFTNND 191
Db      143  ELCPLSFSGFSQMFONCESLT-----SDTSASNVI PHASAIYATTPMLFTNND 191
QY      177  KFSFYNLVSDDGDAIDAKSLTVQGISKLCVQFQNTAAGGACQVVTFSAMANEAPIA 236
Db      177  KFSFYNLVSDDGDAIDAKSLTVQGISKLCVQFQNTAAGGACQVVTFSAMANEAPIA 236
QY      192  SILFQVNRSGAGAAIRGTSTITTENTKSLFPNGSGISNGALTGSAALNINNSAPVI 251
Db      192  SILFQVNRSGAGAAIRGTSTITTENTKSLFPNGSGISNGALTGSAALNINNSAPVI 251
QY      237  FVANVAGVGGIAAVQDGGQGVSSSTSTEDPVVFSFRNTAVFDPGNVAVGGGIYSYGN 296
Db      237  FVANVAGVGGIAAVQDGGQGVSSSTSTEDPVVFSFRNTAVFDPGNVAVGGGIYSYGN 296
QY      252  PSTNATGIYGAIVLT-----GGSLMTS-----GNLSGLVFNNSRSGGAIYANGN 298
Db      252  PSTNATGIYGAIVLT-----GGSLMTS-----GNLSGLVFNNSRSGGAIYANGN 298
QY      297  VAPLNNKTLFLNNVASPV-YIAKQPTSGQASNTSNYDGGGAIKCNQAQAGSNNSGS 355
Db      297  VAPLNNKTLFLNNVASPV-YIAKQPTSGQASNTSNYDGGGAIKCNQAQAGSNNSGS 355
QY      299  VTFENNSDLTFQNTASFNLSLPAFTPTPTTTPAVTP-LLGVGGAIFCTPTPTPTPTGVS 356
Db      299  VTFENNSDLTFQNTASFNLSLPAFTPTPTTTPAVTP-LLGVGGAIFCTPTPTPTPTGVS 356
QY      356  VSPDGEVGFVFPSSNVAAGKGAIAKLVANCGPVQFLRNANDGGAIVLGGSGELSL 415
Db      356  VSPDGEVGFVFPSSNVAAGKGAIAKLVANCGPVQFLRNANDGGAIVLGGSGELSL 415
QY      357  LTISGENSVTFLENIAEQGGAIVGKKISIDSNKSTIPLNGTAGKGAIALPESGELSL 416
Db      357  LTISGENSVTFLENIAEQGGAIVGKKISIDSNKSTIPLNGTAGKGAIALPESGELSL 416
QY      416  ADYGDIIFDGNLKRKTAKENADVNGVTVSSQAI SMGSGGKITTLIRAKAGHQILFNDPIEM 475
Db      416  ADYGDIIFDGNLKRKTAKENADVNGVTVSSQAI SMGSGGKITTLIRAKAGHQILFNDPIEM 475
QY      417  ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKPAITLGAQTGYLYFYDPI-- 465
Db      417  ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKPAITLGAQTGYLYFYDPI-- 465
QY      476  ANGNQPAQSKLKLKIN-----DEGYTGDIIVP-----ANGSTLYQNTIIE 517
Db      476  ANGNQPAQSKLKLKIN-----DEGYTGDIIVP-----ANGSTLYQNTIIE 517
QY      466  TSDLSAASAATAVWVNPKASADG-AYSGTIVFSGETLTATEAPATPATNATSTLNKOLE 524
Db      466  TSDLSAASAATAVWVNPKASADG-AYSGTIVFSGETLTATEAPATPATNATSTLNKOLE 524
QY      518  QGRIVLREKAKLVNLSLQTCGS-LYMEAGSTWDPVTPPQPPQPPAANQLITLSNLHSL 576
Db      518  QGRIVLREKAKLVNLSLQTCGS-LYMEAGSTWDPVTPPQPPQPPAANQLITLSNLHSL 576
QY      525  GGTLLARNGATLNVHNFQDEKSVVIMDAGTT--LATNGANNITDGA---ITLKNLVINL 579
Db      525  GGTLLARNGATLNVHNFQDEKSVVIMDAGTT--LATNGANNITDGA---ITLKNLVINL 579
QY      577  SSLANNAAVTPPTNPPAQDGHPAVIG-STTAGSVTISGPPIFFEDLDITAYDRYDWLGSN 635
Db      577  SSLANNAAVTPPTNPPAQDGHPAVIG-STTAGSVTISGPPIFFEDLDITAYDRYDWLGSN 635
QY      580  DSLDGTAK-----AVVNVQSTNGALTISGTLGLVKNQSDCCDNHGMFKD 624
Db      580  DSLDGTAK-----AVVNVQSTNGALTISGTLGLVKNQSDCCDNHGMFKD 624

; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 455
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0453
; US-09-438-185A-455

Query Match      25.9%; Score 1365.5; DB 2; Length 999;
Best Local Similarity 33.5%; Pred. No. 3.8e-110;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY      636  -QKINVLKQLQTKPPANAPSDL-TLGNEMPKYQGSWKLAWDPTNANGPYTKATWT 693
Db      636  -QKINVLKQLQTKPPANAPSDL-TLGNEMPKYQGSWKLAWDPTNANGPYTKATWT 693
QY      625  LQOVPILELKATSNVTVTTFDLSLGTNGYQOSPYQGTWETIDTTT-----HVTGNWK 679
Db      625  LQOVPILELKATSNVTVTTFDLSLGTNGYQOSPYQGTWETIDTTT-----HVTGNWK 679
QY      694  KTGYNPGPERVASLVPNSILWSILDIRSAHSAIQASVDGRSY-CRGLWVSGVSNFYHNR 752
Db      694  KTGYNPGPERVASLVPNSILWSILDIRSAHSAIQASVDGRSY-CRGLWVSGVSNFYHNR 752
QY      680  KTGYPHPERLAPLIPNSLWANVIDLRAVSOA--SAADGEDVPGKQLSTIGITNFFHANH 737
Db      680  KTGYPHPERLAPLIPNSLWANVIDLRAVSOA--SAADGEDVPGKQLSTIGITNFFHANH 737
QY      753  DALGOGYRISGGYSLGANSY---FGSSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYL 809
Db      753  DALGOGYRISGGYSLGANSY---FGSSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYL 809
QY      738  TGDARSYRHMGGYLI--NTYTRIIPDAALSIGFQQLFTKSKDYLVGCHSNVYFATVYS 795
Db      738  TGDARSYRHMGGYLI--NTYTRIIPDAALSIGFQQLFTKSKDYLVGCHSNVYFATVYS 795
QY      810  STQOALCG-SYLFDAFIRASVFGNQHMKTSYTPAESDVWRDNNCLAGETGAGLPIVI 868
Db      810  STQOALCG-SYLFDAFIRASVFGNQHMKTSYTPAESDVWRDNNCLAGETGAGLPIVI 868
QY      796  NITKSLFGSSRFFSGTSRVTYSRNEKVKTSYTKLPKGRCSWNNCMLGEGLNLPITL 855
Db      796  NITKSLFGSSRFFSGTSRVTYSRNEKVKTSYTKLPKGRCSWNNCMLGEGLNLPITL 855
QY      869  TSPKLYLNELRPPVQAEFSYADHESFTREGDOARAFKSHLNLNLSVPVCVKDRCSSHP 928
Db      869  TSPKLYLNELRPPVQAEFSYADHESFTREGDOARAFKSHLNLNLSVPVCVKDRCSSHP 928
QY      856  SSRILNLKQIIPFVKAEVAYATHGGIQENTPEGRIFGHGHLNVAVPVGRFGKSHNRP 915
Db      856  SSRILNLKQIIPFVKAEVAYATHGGIQENTPEGRIFGHGHLNVAVPVGRFGKSHNRP 915
QY      929  NKYSFMAAYICDAYRTISGTETTLTSHOBTMTTDAFHLARHGWWVRGSMYASLTSNIEVY 988
Db      929  NKYSFMAAYICDAYRTISGTETTLTSHOBTMTTDAFHLARHGWWVRGSMYASLTSNIEVY 988
QY      916  DFYTIIVAYADPYRHRNPDCDITLIPNGATWTSIGNLSTRLLVQASSTSVNDVLEIF 975
Db      916  DFYTIIVAYADPYRHRNPDCDITLIPNGATWTSIGNLSTRLLVQASSTSVNDVLEIF 975
QY      989  GHGRVEYRDASRGYGLSAGSRVRF 1012
Db      989  GHGRVEYRDASRGYGLSAGSRVRF 1012
QY      976  GHGCDIRTSRQYTLIDIGSKLRF 999
Db      976  GHGCDIRTSRQYTLIDIGSKLRF 999

RESULT 22
US-09-430-723-2
; Sequence 2, Application US/09430723
; Patent No. 6607730
; GENERAL INFORMATION:
; APPLICANT: Mordin, Andrew
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: US95 THEREOF
; FILE REFERENCE: 19721-011
; CURRENT APPLICATION NUMBER: US/09/430,723
; CURRENT FILING DATE: 1999-10-29
; EARLIER APPLICATION NUMBER: 60/106,590
; EARLIER FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/133,071
; EARLIER FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-430-723-2

Query Match      25.8%; Score 1359.5; DB 2; Length 973;
Best Local Similarity 33.4%; Pred. No. 1.2e-109;
Matches 349; Conservative 171; Mismatches 421; Indels 103; Gaps 27;

QY      1  MOTSFHKFPLSMILAYSCCSLNGGVAABIMVQGIYDGETLTVSPYTVIGPSPGTVTF 60
Db      1  MTSIRKFLISTTLAPCFAST---AFTVEVIMSENFDDGSSGKI-PPYTTLSDFRGLTLCI 56
QY      61  SAGELTKNLDNSIAALPLSCFNLGSGFTVLGRGHSLTPENIRTSNNGAALSNAADG- 119
Db      61  SAGELTKNLDNSIAALPLSCFNLGSGFTVLGRGHSLTPENIRTSNNGAALSNAADG- 119
QY      57  FSGDLIYANLDNAISRTSSCFNRRAGALQILKGKGVFSLNIRSSADGAIAISSVITQNP 116
Db      57  FSGDLIYANLDNAISRTSSCFNRRAGALQILKGKGVFSLNIRSSADGAIAISSVITQNP 116
QY      120  ---LFTIEGKELSPFNCNSLLAVLPAATNKGSOPTTTTSTPSNGTIYKTDLLLNNE 176
Db      120  ---LFTIEGKELSPFNCNSLLAVLPAATNKGSOPTTTTSTPSNGTIYKTDLLLNNE 176
QY      117  ELCPLSFSGFSQMFONCESLT-----SDTSASNVI PHASAIYATTPMLFTNND 165
Db      117  ELCPLSFSGFSQMFONCESLT-----SDTSASNVI PHASAIYATTPMLFTNND 165
QY      177  KFSFYNLVSDDGDAIDAKSLTVQGISKLCVQFQNTAAGGACQVVTFSAMANEAPIA 236
Db      177  KFSFYNLVSDDGDAIDAKSLTVQGISKLCVQFQNTAAGGACQVVTFSAMANEAPIA 236
QY      166  SILFQVNRSGAGAAIRGTSTITTENTKSLFPNGSGISNGALTGSAALNINNSAPVI 225
Db      166  SILFQVNRSGAGAAIRGTSTITTENTKSLFPNGSGISNGALTGSAALNINNSAPVI 225
QY      237  FVANVAGVGGIAAVQDGGQGVSSSTSTEDPVVFSFRNTAVFDPGNVAVGGGIYSYGN 296
Db      237  FVANVAGVGGIAAVQDGGQGVSSSTSTEDPVVFSFRNTAVFDPGNVAVGGGIYSYGN 296
```

```
Db      226  PSTNATGYGGAIYIT-----GCSMLTS-----GNLSGVLFVNNSRSGGAIYANGN 272
Qy      237  VAFNNGKTLFLNNVASPV-YIAAKOPTSGQASNTSNNGDGGAIFCKNGAQAGSNNSGS 355
Db      273  VTFSSNNDLTFQNTASQNSLPAPTPTPPPAVTP-LLGYGGAIFCTPPATPPPTGV-S 330
Qy      356  VSPDEGVVFFSSNVAAGKGGAIYAKKLSVANGCPVQFLRNITANDGGAIYIGESGELS 415
Db      331  LTISGENSVTFLENTASQGGALYKKTISIDSNKSTIFLGNITAGKGAIAIPESGELS 390
Qy      416  ADYGDIIFDGNLKRTAKENAAVDVNGVTVSSQAI SMGSGKLTIRAKAGHOILFNDDPIEM 475
Db      391  ANQGDILFNKNLSITSG-----TPTRNSIHFGKDAKAPATLGNTOGYTLYFYFDPI-- 439
Qy      476  ANGNNOQAQSSKLLKIN-----DGBGYTGDIYVF-----ANGSSTLYQNVITIE 517
Db      440  TSDDLASAASAAATVVVNPKNASADG-AYSGTIVFSGETLTATEAATPATATSTLNQKLE 498
Qy      518  QQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQPPQPPAANQLITLSNHL 576
Db      499  GGTALRNGATLNVNFTQDEKSVIMDAGTT--LATNGANNITDGA---ITLANKLVINL 553
Qy      577  SSLANNAVTNPPTPPAQDSHPAIG-STTAGSVTISGPIFFEDLDDTAYDRVDMG 635
Db      554  DSLDGTKA-----AVNVQSTNGALUTISGTLGVKNQDCCDNHGMFNKD 598
Qy      636  -QKINVLKLQGTKPPANAPDL-TLGNEMPKYGYQGSWKLAWDNTANNGPYTLKATWT 693
Db      599  LQQVFILELKATSNVTITDFSLGTNGYQOSPYGQGTWBEFTIDTT-----HVTGNWK 653
Qy      694  KTYNPGPERVASLVPNSLWGSILDIRSAHSAIAQASVDGRSY-CRGLWVSGVSNPFYHDR 752
Db      654  KTYLPHPERLAPLNSLWANNVILDRAVSQ--SAADGEDVPKQLSITGITNPFHANH 711
Qy      753  DALGGYVYISGYSGLANSY---FGSSMFGLAFTVEVGRSKDYVVCRSNHACIGSYVL 809
Db      712  TGDARSYRHMGGYILI--NTYTRITPDAAISLGFQQLFTKSKDYLVHGHNSVVPATVYS 769
Qy      810  STQQAALCG-SYLFGDFAIRASYGFGNQHMKTSTYTPAESDVRWNNCLAGIGAGLPVI 868
Db      770  NITKSLFGSSRFSGTSRVTVYSRNEKVTSTYKLPKGRCSWNNCWLEGNLPTL 829
Qy      869  TPSKYLINELRPFVQARSYADHESFTREGDOARAPKSGHLINLSVPVGVKPDRCSTHP 928
Db      830  SSRILNLKQIIPFVKAEVAYATHGQIQENTPBGRIFGHGLILNVAVPVGVFRGKNSHRP 889
Qy      929  NKYSFMAAVIDAYRTISGTETTLSSHQETWTTDAFLARHGCVVVRGSMYASLTSNIEVY 988
Db      890  DFTIIVAYAPDVYRHNPDCTTLTPINGATWTSIGNNLTRSLLVQASSHTSVNDVLEIF 949
Qy      989  GHGREYDRDASRGYGLSAGSRVRF 1012
Db      950  GHGCGDIRTSRKYTLIDIGSKLRF 973
```

## RESULT 23

```
US-09-428-122-2
; Sequence 2, Application US/09428122
; Patent No. 6872814
; GENERAL INFORMATION:
; APPLICANT: Connaught Laboratories Limited
; APPLICANT: Murdin et al.
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 19721-007-019
; CURRENT APPLICATION NUMBER: US/09/428,122
; CURRENT FILING DATE: 1999-10-27
; EARLIER APPLICATION NUMBER: 60/106,046
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/132,271
; EARLIER FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 4
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-428-122-2
```

```
Query Match      21.4%; Score 1128; DB 2; Length 928;
Best Local Similarity 29.2%; Pred. No. 2.2e-89;
Matches 311; Conservative 169; Mismatches 394; Indels 192; Gaps 32;
```

```
Qy      1  MQTSPHFHFFLSMILAYSCSLNGGGYAAE-TWVPGIYDGETLTVSPFPYTVIGDPSPVTV 59
Db      1  MKSFPFPKPVFTFAIFPLSMI-----ATETVLDSSASFDGNK-NGNFVSRSSQEDAGTYY 54
Qy      60  PSAGELTKLNLDNIAALPLSCFQNLGSGFTVLGRGHSITPENIRTST-NGAALSNSAAD 118
Db      55  LFKGNVTLLENIPGVTGTAITKSCFNNTKGDLTFTGNGNSLLFTQVDAGTVAGA VSSVVD 114
Qy      119  GLFTIEGFKELSPNCNSLLAVLPAAATNKGSOPTTTSTPSPNGTIYKTDLLLLNNEKF 178
Db      115  KSTTFIGFSSLSF-----IASPGSSITTKGAVSCSTGS-----LSLTKNVSL 157
Qy      179  SFYNLSYSGDGAIDAKSLTVQGISKLCVFOENTAQAQAGGACQVVTSPSAMANEAPAFV 238
Db      158  LFSKNFSTDNGGAI TAKTLTSLTGTMSALFSENSTSKKGAIQTSDALITGNQGEVSF- 216
Qy      239  ANVAGVRGGGIAAVQDQOQGVSSSTSTEDPVVSPSRNTAVFDCGNVARVGGIYSYGNVA 298
Db      217  -----SDNTSSDS-----GAAIFTEASVT 235
Qy      299  FLNNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGDGGAIFCKNGAQAGSNNSGVSF 358
Db      236  ISNNAKVSFIDN-----KVTGASSSTTGDG-SGCAICAYK-----TSTDTKVTL 278
Qy      359  DGEGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQFLRNIA-----DGGAIYIGESGELS 413
Db      279  TGNQMLLFSNNTSTTAGGAIYVKKLELASGGLTLFSSNVNGGTAPKGGAI AIEDSGELS 338
Qy      414  LSADYDGIIFDGNLKRTAKENAAVDVNGVTVSSQAI SMGSGKLTIRAKAGHOILFN 473
Db      339  LSADSGIIVFLGN---TVTSTTPGTN-----RSSIDIGTSKMTALSAAGRAIYFDP 390
Qy      474  EMANGNNOQAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNV 515
Db      391  TTGSSTT-----VTDVLKVNETPADSALQYTGNIIFTGKELSETEAASKNLTSLKLPVT 446
Qy      516  IEQGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQPPQPPAANQLITLSNHL 574
Db      447  LSGGTLSLKHGVTLTQTAQQADSRLEMDVGTTL-----PADTS--TINNLI 494
Qy      575  SLSSLANNAVTNPPTPPAQDSHPAIGS-TTAGSVTISGPIFFEDLDDTAYDRYDMLG 633
Db      495  NISSI-----DGAKKAKETKATSKNLTLSGTITLLDPTGTFYENHS-IR 538
Qy      634  SNQKINVLKQL-CTKPPANAPSDLTIGNEMPKYGYQGSW-KLAWDPNTANNGPYTLKAT 691
Db      539  NPQSYDILELKASGTVTSTAVTPDPIMGKFP-HYGYQGTWGPVWGTGASTATF---N 593
Qy      692  WTKTYNPGPERVASLVPNSLWGSILDIRSAHSAIAQASVDGRSYCRGLWVSGVSNPFVHD 751
Db      594  WTKTYGYPNPERIGSLVPSLWNAFIDISLHYLMETANEBLQGDRAFCWAGLSNFPFKD 653
Qy      752  RDALGQGVYISGYSGLANSYFGS-SMFGLAFTVEVGRSKDYVVCRSNHACIGSYVLS 810
Db      654  STKTRGFRHLSGGYVIGGNLHTCSDKLSAFCQLFGRDREDYFVAKNQ-----GTVYGG 708
Qy      811  TQQAALCGSYLFGDAFI-----RASYGFGNQHMKTSTYTFABE 846
Db      709  TLY-----YQHNETYISLPCKLRPCSLSYVPTETPVLFSGNLSYTHTDNDLTKYTYPT 763
Qy      847  SDVWMDNCLAGEIAGLPVITPSKLYLRELRFVQAEFSYADHESFTREGDOARAPKS 906
```

Db	764	VKSGWNSDFALEFGGRAPICIDESALP-EQYMPFMKLFQVVAHQEKFQCKQTEAREFGS	822
Qy	907	GHLNLVSVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETILLSHQETTTDAPHL	966
Db	823	SRLVNLALPIGIRFDKESDCQDATYNTLGTGYVDLVRSPNDCPTTTLIRISGDSWKTFGTNL	882
Qy	967	ARHGUVVRGSMVASLTSNLEVYGHGREYEDASRGYGLSAGSRVRF	1012
Db	883	ARQALVLRAGNHFCFNSNFEPASQFSFELRGSSRNYNVDLGAQYQF	928
RESULT 24			
US-09-198-452A-478			
; Sequence 478, Application US/09198452A			
; Patent No. 6559294			
; GENERAL INFORMATION:			
; APPLICANT: Griffiths, R.			
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection			
; FILE REFERENCE: 9710-003-999			
; CURRENT FILING DATE: 1998-11-24			
; NUMBER OF SEQ ID NOS: 6849			
; SEQ ID NO 478			
; LENGTH: 949			
; TYPE: PRT			
; ORGANISM: Chlamydia pneumoniae			
; FEATURE:			
; NAME/KEY: SITE			
; LOCATION: 1...949			
; OTHER INFORMATION: Xaa=unknown or other			
US-09-198-452A-478			
Query Match			
Best Local Similarity 20.7%; Score 1090; DB 2; Length 949;			
Best Local Similarity 28.8%; Pred. No. 5e-86;			
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;			
Qy	1	MQTSFHKFFLSMILAYSC--CSLNGGGYAAEIMVPGQIYDGETLTVSF-----PYTVI	51
Db	22	MKTSIPMWLVSVLAFSCHLQSLAN----BELLSPDSDSFGNIDSGTFTPKTSATTYSLT	77
Qy	52	GDPSGTTVSAGELTLKNLNSIAALPL--SCFNGLLGSFTVLGRGSHLFEINRTSNG	109
Db	78	GD---VFEPYRKGKT-----PLSDSCFKQTQNTDLNLTFLNGHSLTFGFDAGTHA	123
Qy	110	AALNSAADGLFTIEGKELSFNSNCNSLLAVLPATTNKGSQTPTTSTPSNGTIYSKTD	169
Db	124	GAASATANKNLTFSGFSLISFDS-----SPSTTVTTGGQTLSSAGG	165
Qy	170	LLILNNEKFSYSLNLSGDGAIDAKSLTVQGISKLCVFOENTAAQADGGACQVVTSPSAM	229
Db	166	VLENIRKLVVAGNFSTADGAIGASFLLTGTSGDALFNSNSSTTKGA-----	215
Qy	230	ANEAPIAFVANVAGRVGGIAAQQDGGQGVSSSTSTEDPVVPSRNT-AVEFDGNVARVG	288
Db	216	-----IATTAGAR-----IANNTGXRVFLSNIASTS	241
Qy	289	GGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNYGGGAIKCKGAQA	348
Db	242	CGAIDDEGTSILSNKKFLYP-----EGNAAKTT-----GGAI-CNTKASG	280
Qy	349	G-----SNNSGSVPDGBGVVFFSVNVAAGKGAIAKLSVANCGVPQFLRNLIAND---	400
Db	281	SPELIISNKK-----TLIFASNVAEFTSGGAHAKKLALSSGGFTFLRNVSATP	331
Qy	401	GGAIYLGESGELSADYDGIIFDGNLKRXTAKENADVNGVTSSQAIMSGSGKITTLL	459
Db	332	KGGAISIDASGELSLSAETGNITFVRNLTLLT--GSTD-----TPKRNAINIGSNGKFTFL	385
Qy	460	RAKAGHQILFNDPIEMANGNNQPAQSKLLKINDGEG-----YTGDIVFA-----	504
Db	386	RAAKNHTTIFFYDPI-----TSEGTSSDVLKINNGSAGALNPYQGTILFSGETTLTADBLK	439

RESULT 25

US-09-438-185A-453

; Sequence 453, Application US/09438185A

; Patent No. 6822071

; GENERAL INFORMATION:

; APPLICANT: Stephens, Richard

; APPLICANT: Mitchell, Wayne

; APPLICANT: Kalman, Sue

; APPLICANT: Davis, Ronald

; APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence

; FILE REFERENCE: 018941-000411US

; CURRENT APPLICATION NUMBER: US/09/438,185A

; CURRENT FILING DATE: 2002-03-13

; PRIOR APPLICATION NUMBER: US 60/108,279

; PRIOR FILING DATE: 1998-11-12

; PRIOR APPLICATION NUMBER: US 60/128,606

; PRIOR FILING DATE: 1999-04-08

; NUMBER OF SEQ ID NOS: 1074

; SOFTWARE: FASTSEQ for Windows Version 3.0

; SEQ ID NO 453

; LENGTH: 967

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

; FEATURE:

; OTHER INFORMATION: CPn0451

US-09-438-185A-453

Query Match 20.7%; Score 1090; DB 2; Length 967;

Best Local Similarity 28.8%; Pred. No. 5.2e-86;

Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;





Db	533	-----SHTIN-----QAFTPLVVPFAATAASDIYDALLTSPVQTPPEHYGYQGWHEAT	582
Qy	676	W-DPNTANNPPTLKATWTKTGYNPGPERVASLVNSLWMSILDIRSAHSAIQASVDGRS	734
Db	583	WADTSTAKSG---TMTWTTGYNPNPERRASVVPDLSLWASFTDITLQQIMTSQANSIY	638
Qy	735	YCRGLWVGSVSNFFYHDDRDLAQGVRYISGGYSLGANSY-FGSSMFGLAFTVEFGRSKDY	793
Db	639	QQRGLWASGTANFFHKDKSGTQNAFHKHSYGYVGSADDFSENIFSVAFQCFKDKDL	698
Qy	794	VVCRSNHACIGSVLSTQOALCGSYL--FGDAF-----IRASYGFGNQHMTS	840
Db	699	FIVENTSHNYLASLYLQHRFLGGLPMPSPGSTMDMLKDIPLILNAQLSYSTKNDMDTR	758
Qy	841	YTPAESDVRWNNCLAGIBAGLPVITPFSKLYLNLPPFVQAEFSYADHESFTREGDQ	900
Db	759	YTSYPEAQGSWTNNSGALBELGSLALYLKPEAPPFQGYFPLKQAVYSRQONFKESGAE	818
Qy	901	ARAPKSHLLNLVSPVGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETILLSHOBTWT	960
Db	819	ARAFDDGDLVNCISIPVGIRLEKISEDEKNFEISLAYIGDVYRKNPRSRISLTVSGASWT	878
Qy	961	TDAFLHARGVVRGSMVYASLTNSIEVYGHGRYERDASRGYGLSAGSRVF	1012
Db	879	SICKNLARQAFTLASAGSHLTLSPHVELSGEAYELGSAHIYNVDCGLRYSF	930
RESULT 27			
US-09-438-185A-448			
; Sequence 448, Application US/09438185A			
; Patent No. 6822071			
; GENERAL INFORMATION:			
; APPLICANT: Stephens, Richard			
; APPLICANT: Mitchell, Wayne			
; APPLICANT: Kalman, Sue			
; APPLICANT: Davis, Ronald			
; APPLICANT: The Regents of the University of California			
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence			
; FILE REFERENCE: 018941-000411US			
; CURRENT APPLICATION NUMBER: US/09/438,185A			
; CURRENT FILING DATE: 2002-03-13			
; PRIOR APPLICATION NUMBER: US 60/108,279			
; PRIOR FILING DATE: 1998-11-12			
; PRIOR APPLICATION NUMBER: US 60/128,606			
; PRIOR FILING DATE: 1999-04-08			
; NUMBER OF SEQ ID NOS: 1074			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 448			
; LENGTH: 938			
; TYPE: PRT			
; ORGANISM: Chlamydia pneumoniae			
; FEATURE:			
; OTHER INFORMATION: CFP0446			
US-09-438-185A-448			
Query Match 20.5%; Score 1079; DB 2; Length 938;			
Best Local Similarity 28.3%; Pred. No. 4.5e-85;			
Matches 303; Conservative 160; Mismatches 407; Indels 202; Gaps 29;			
Qy	1	MOTSFKFPFLSMILAYSCCSLNGGYAAEI-MVPGQYDGETLTVSPYTVIGDPSGTV	59
Db	9	MKIPLHKLISSTLV-TPILISIAIYAGDASLSPDSDGAGGSTFTPKST-ADANGTNY	66
Qy	60	FSAGELTKNLDNSIAALPLSCFNLGLSPVLGRGSLTFENIRTSNAGALSNSAADG	119
Db	67	VLSGNVINDAGK-TALTGCCPTTETGDLTFKGGYSPFNTVDAGSNAGAAATYADK	125
Qy	120	LFTIEGFKELSFNSCNLSLAVLPAAITNKGSTPTTT---STPSNGTYIKTDLLLNNE	176
Db	126	ALTGTGNSLSF-----INAPGTWASGKSTLSSAGALNLTGNTILFQNV-----	172
Qy	177	KFSFYSNLVSGDGAIDAKSLTVQGISKLVPQENTAQDAGGACQVWTSFSMANEAPIA	236

Db	173	-----SNEANNNGAIITTKTSLISGNTSSITFTNSAKKLGAIYSSAAASISGNTGQLV	227
Qy	237	FVANVAGVYRGGAIAAVODGQGVSSSTSTEDVPVVSFRNTAVEPDGNVARVUGGGIYSYGN	296
Db	228	FMNN-KGTGGALGFE-----ASSSTIQNSLFFSGNTATD-----	263
Qy	297	VAFLLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNNTYDGGGAIFCKGGAQAGSNNSGV	356
Db	264	-----AA-----GKGAIYCEK-----TGETPTL	282
Qy	357	SPDGGVVPFSSNVAAGGGAIIYAKKLSVANGCPVQFLN-----IANDGGAIYLGESGE	411
Db	283	TTISGKSLTFAENSSVTOGGAICAHGLDLSAAGPTLFSNNRCNTAAGKGAIAIADSGS	342
Qy	412	LSLSADYGDIIIFDGNLKRTAKENAADVNGVTVSSQAISMGSGGKITTIRAKAGHQLFND	471
Db	343	LSLSANGDITPLGN-----TLTSTAPTSTRNRIYLGSSAKITNLEAAQOSIYFYD	395
Qy	472	PIEMANGN-----NPPAQSSKLLKINDGEGYTDIVPA-----NGSSTLY	511
Db	396	PIASNTTGASDVLTINQPDNSPL-----DYSGTIVFSGEKLSDAKAADNFTSILK	448
Qy	512	QNVTTIEQGRIVLREKAKLSVNSLSOTGGSLYNEAGSTWDTVPQPPQPPAANQLITLN	571
Db	449	QPLALASGTLALKGNVELDVNGFTQTEGSTLL-----MQPGTKLKADTEAISLTK	498
Qy	572	LHLSLSLLANNVATNPPTNPPAQDSHPAIGVSTTAG-----SVTISGPIFFEDLDDTAYDR	628
Db	499	LVVDLSALEGNKSVS-----IFTAGANKTILTSPLVFPQSSGNFYE-540	
Qy	629	YDWLGSNQKINVLKLQGTGP-----PANAPSDLTIG-----NEMPKYQGSWKLA	675
Db	541	-----SHTIN-----QAFTPLVVPFAATAASDIYDALLTSPVQTPPEHYGYQGWHEAT	590
Qy	676	W-DPNTANNPPTLKATWTKTGYNPGPERVASLVNSLWMSILDIRSAHSAIQASVDGRS	734
Db	591	WADTSTAKSG---TMTWTTGYNPNPERRASVVPDLSLWASFTDITLQQIMTSQANSIY	646
Qy	735	YCRGLWVGSVSNFFYHDDRDLAQGVRYISGGYSLGANSY-FGSSMFGLAFTVEFGRSKDY	793
Db	647	QQRGLWASGTANFFHKDKSGTQNAFHKHSYGYVGSADDFSENIFSVAFQCFKDKDL	706
Qy	794	VVCRSNHACIGSVLSTQOALCGSYL--FGDAF-----IRASYGFGNQHMTS	840
Db	707	FIVENTSHNYLASLYLQHRFLGGLPMPSPGSTMDMLKDIPLILNAQLSYSTKNDMDTR	766
Qy	841	YTPAESDVRWNNCLAGIBAGLPVITPFSKLYLNLPPFVQAEFSYADHESFTREGDQ	900
Db	767	YTSYPEAQGSWTNNSGALBELGSLALYLKPEAPPFQGYFPLKQAVYSRQONFKESGAE	826
Qy	901	ARAPKSHLLNLVSPVGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETILLSHOBTWT	960
Db	827	ARAFDDGDLVNCISIPVGIRLEKISEDEKNFEISLAYIGDVYRKNPRSRISLTVSGASWT	886
Qy	961	TDAFLHARGVVRGSMVYASLTNSIEVYGHGRYERDASRGYGLSAGSRVF	1012
Db	887	SICKNLARQAFTLASAGSHLTLSPHVELSGEAYELGSAHIYNVDCGLRYSF	938

RESULT 28  
US-09-438-185A-447  
; Sequence 447, Application US/09438185A  
; Patent No. 6822071  
; GENERAL INFORMATION:  
; APPLICANT: Stephens, Richard  
; APPLICANT: Mitchell, Wayne  
; APPLICANT: Kalman, Sue  
; APPLICANT: Davis, Ronald  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence  
; FILE REFERENCE: 018941-000411US  
; CURRENT APPLICATION NUMBER: US/09/438,185A

```
/ CURRENT FILING DATE: 2002-03-13
/ PRIOR APPLICATION NUMBER: US 60/108,279
/ PRIOR FILING DATE: 1998-11-12
/ PRIOR APPLICATION NUMBER: US 60/128,606
/ PRIOR FILING DATE: 1999-04-08
/ NUMBER OF SEQ ID NOS: 1074
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 447
/ LENGTH: 947
/ TYPE: PRN
/ ORGANISM: Chlamydia pneumoniae
/ FEATURE:
/ OTHER INFORMATION: Cpn0445
US-09-438-185A-447

Query Match      19.8%; Score 1041; DB 2; Length 947;
Best Local Similarity 29.0%; Pred. No. 9.9e-82;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

Qy 1 MQTSFHKPFLSMILAYSCSLNGGGYAAEIMV--PQGIYDGETLTVSPFYTVIGDPGTT 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12 MKSSVSWLFFSSIPFLSLSI-----VAAEVTLDSNNSYDGSNGTTFVFTTDAAGTT 67

Qy 59 VPSAGELTLKNDNSIAALPLSCFNGLLGSFTVLGRGHSITPENIRT-STNGAALNSAA 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 YLLSDVDFQNGALGIPLSCGFLEAGDGLTPQNGHAKLPAPINAGSSAGTAVASTSAA 127

Qy 118 DGLFTIEGFKELSFNCNLSLAVLPAATNKGSOPTTTTSPNSGTYISKTDLALLNEX 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 DNKLNFDFSLRSLIISCPSLLL-----SPT-----GQCALKSVGNLSUTGNSQ 170

Qy 178 SPFYENLVGDCGADDAKSLTVQIGSKLCVPOENTAOAGGACQVVTFSAMANEAPAP 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 IIFTQNFSDNGGVINTKFLSGTSQFASFRNOA-----206

Qy 238 VANVAGVGGGTAAGVQDQGGVSSSTEDP-VVFSRNTAVEPDGNVARVGGGIYSYGN 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 ---FTGKGQGVYA-----TGTTIENSPOIVSFQNLA---KGS---GGALYSTDN 249

Qy 297 VAFLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNTYDGGAIFCNGAQAGSNNSGV 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 CSITDNFQVIFDNGSA---WEAAQ-----AQQGAICC-----TTTDTKTV 285

Qy 357 SPDGEVGVFFSSNVAAGGAIYAKLKVANGCPVQFLNTAND-----GGAIYLGESG 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 TLTGKKNLSPTNNTALTYGGAISGLKVSISAGGPTLFQSNISGSSAGGQGGGAINASAG 345

Qy 411 ELSLSADYGDIDFQGNLKRKTAKENAAVNGVTVSSQAIWSGSGGKITTLRAKAGHOILFN 470
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 ELALSATSGDITFNN-----QVTNGSTSTRNAINIDTAKVTSIRATGQSIYFY 396

Qy 471 DPEIANGNNOQAQSSKLLKINDRG-----YTGDIVF-----ANGSSTLYQ 512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 DPIT---NPGTAASDTDLNLMLDANGSEIEYGGAI VFSGEKLSPTKEAIAANVTSTIRQ 452

Qy 513 NVTIBQGRIVLRBAKLVSNSLSQTGGS-LYMEAGSTWDFVTPQPQPPAANQILITLNS 571
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 453 PAVLARGDLVLRDGVTVTFKDLTQSPGRILMDGGTTLS-----AKEANLSLNG 501

Qy 572 LHLSSLSLANNNAVTPNPPNPAQDSHAPVIGSTTAGSVTISGRIFFPFDLDDTAYDRYDM 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 LAVNLSSLDGTNKA-----ALKTEAADKNISLSGTTALIDTREGSFYENNN- 546

Qy 632 LGSNOKINVLKQLQTKPPAN-----APSDITLGNEMPKYGYQGSWKLAMPDNTANNGP 685
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 LKASATYPLLEL---TTAGANGTITLGAISTLTQLQEPETHYQGNWQLSWANATSSK-- 601

Qy 686 YLKATWTYKYNCPGPVRSVLVPSNLWGLSILDIRSAHSIAQASVDGRSVCRLWVGVS 745
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 -IGSINWTRTGYIPSPERKSNLPLNSLWGNFDIRSINQLIETKSSGPPFERELWSGIA 660

Qy 746 NFFYHRRDALGQGYRYISGGYSLGANSYP-GSSMFGLAFTVEFGSKDYVVCRSNHHACI 804
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 241 -SPISNKAISFINN-----SVTATSAT-----CGAIYC-----SSTSAPKPV 277  
Qy 356 VSFDEGEVVPFSNVAAGKGAIIYAKKLSVANGCPVOFLRNTAND-----GGAIVLGESG 410  
Db 278 L7LSDNGELNFIQNTAITSGGAIYTDNLVSSGGTFLFKNGSAIDTAAPLGGAIADSG 337  
Qy 411 ELSLSADYGDIIIFDGNLKTAKENADVNGVTVSSQAIISMG--SGGKITTLRAKAGHOILF 469  
Db 338 SLSLALGDDITFEQN-----TVVKAS--SSQTTTNSINIGTNNAKIVQLRASQGTIYF 392  
Qy 470 NDIPIEMANGNNOPAOSSKLLKINDGE-----GYTGDIVFA-----NGSSTLY 511  
Db 393 YDPIT-----TSITAALSDALNGLPDLAGNPAYQGTIVFSGEKLSEAEAEADNLKSTIQ 448  
Qy 512 QNVTIEQGRIVIREKAKLSVNSLSQTGGS--LYNEAGSTWDFVTPPQPQPPAANQLITLS 570  
Db 449 QPLTLAGGQSLKSGVTLVAKSPSQSGTLLMDAGTTLETADG-----ITIN 496  
Qy 571 NLHLSLSSLLANNAVTPNPPAODSHPAVIGSTTAG--SVTISGPIFFEDLDDTAYDRY 629  
Db 497 NLVLNVDL-----KETKATLKATQASQTVTLSSLSLVDPSGNVYEDV 541  
Qy 630 DMLGSNOKINVLKQLGKTKPPANAP--SDLT---LGNEMPKYGYQGSWKLAWDPTNANGP 685  
Db 542 SW--NNPOV--FSCUTLTADDPANIHITDLAADPLEKNPIHWGYQGNWALSQWEDTATKSK 598  
Qy 686 YTLKATWTKTGNPGRVAVSLVPSNLWSGILDIRSAHSAIQASVDGRSYCRGLWVGVS 745  
Db 599 -AATLTWTKTGNPNPERRGTLVANTLWGSFVDVRSIQQLVATKVRSQOETRGICEGIS 657  
Qy 746 NFPHDRDALCGQYRISIGSYISLGANSYFGS--SMFGLAFTVEVGRSKDYVVCRSNHHACI 804  
Db 658 NFPHKDSKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKORDHPINKRASAYA 717  
Qy 805 GSVYLSQQALCG---SYLFG-----DAFIRASYGFGNQHKMTSYTFAEBSDVVR 851  
Db 718 ASLHLOHLATLSPSLRYLPGSEBQVLFDAQI--SYIYSKNTMKTYTQAPKGESSW 775  
Qy 852 DNNCLAGEIGAGLPVITPISKLYLNLRLPFVQAEFSYADHESFTEBGDQ--ARAFKSGHLL 910  
Db 776 YNDGCALELASSLPHTALSHEGLPHAYFPFIKVEASYIHQDSPKERNITLVRSPDSGLI 835  
Qy 911 NLSVPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAFHLARHG 970  
Db 836 NVSPVIGITFERFSRNERASYEATVIYVADVVRKNPDCCTTALLINNTSWKTGTNLRSQA 895  
Qy 971 VVRGSMYASLTSNIEVYGHGRVEYRDASRGYGLSAGSRVRF 1012  
Db 896 GIGRAGIFYAFSPNLEVTLSMEIRGSSRSYNADLGKQFP 937

RESULT 30

US-09-198-452A-472  
; Sequence 472, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Grifflais, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
; TITLE OF INVENTION: and treatment of infection  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 472  
; LENGTH: 927  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 1...927  
; OTHER INFORMATION: Xaa=unknown or other  
US-09-198-452A-472

Query Match 18.9%; Score 994.5; DB 2; Length 927;  
Best Local Similarity 29.0%; Pred. No. 1.1e-77;  
Matches 308; Conservative 157; Mismatches 413; Indels 183; Gaps 38;  
Qy 1 MOTSPHKFPLSMILAYSCSLGGGYAAEIMV---PQGIYDGE--TLTVSFPFVTVIGDPSG 56  
Db 1 MKSSLHWFLLISSSLALP--LSLNFSAFAAVWEINLGPFTNSFGPGTYT---PPAQTNDAG 56  
Qy 57 TTVPSAGELTLKNDLNSSTAALPLSCFGNLGSFTVLGRGHSFLTPEIRITSTNGAALSNSA 116  
Db 57 TYINTUTGDSYINA--GSPALTASCPEKTTGNLSFGHGQYQLLQNLIDAGAN--CIFTNTA 114  
Qy 117 ADGLFTIEGFKELFSNCSNLSLAVLPAATTKNGSQTPTTTSTPSNGTIYKTDLLLNNE 176  
Db 115 ANKLASFSGFSYLSL-----IQTNTATGTGAIKSTGACSQSNY 154  
Qy 177 KPSFVSNLVSGDGAIDAKSLTVQIGISKLCVQENTAQADGACQVTVTSFSAMANEAPIA 236  
Db 155 SCYFQGNFSDNGGALQSSISL--SLNPNLTFAKNKATQKGGAL-----YST----- 200  
Qy 237 FVANVAGVRGGGIAAVQDQCGVSSSTSTEDPVWFSRNTAVEFDGCVARVGGIYSYGN 296  
Db 201 -----GGIT-----INNTLNS-----ASFSENTA-----ANNGAIYTEAS 231  
Qy 297 VAPLNNKCTL--FLNNVASPVYIAAKQPTSGQASNTSNNGDGGAIKCKGAQAGSNNGS 355  
Db 232 -SPISNKAISFINN-----SVTATSAT-----GGAIVC-----SSTSAPKPV 268  
Qy 356 VSFDEGEVVPFSNVAAGKGAIIYAKKLSVANGCPVOFLRNTAND-----GGAIVLGESG 410  
Db 269 L7LSDNGELNFIQNTAITSGGAIYTDNLVSSGGTFLFKNGSAIDTAAPLGGAIADSG 328  
Qy 411 ELSLSADYGDIIIFDGNLKTAKENADVNGVTVSSQAIISMG--SGGKITTLRAKAGHOILF 469  
Db 329 SLSLALGDDITFEQN-----TVVKAS--SSQTTTNSINIGTNNAKIVQLRASQGTIYF 393  
Qy 470 NDIPIEMANGNNOPAOSSKLLKINDGE-----GYTGDIVFA-----NGSSTLY 511  
Db 384 YDPIT-----TSITAALSDALNGLPDLAGNPAYQGTIVFSGEKLSEAEAEADNLKSTIQ 439  
Qy 512 QNVTIEQGRIVIREKAKLSVNSLSQTGGS--LYNEAGSTWDFVTPPQPQPPAANQLITLS 570  
Db 440 QPLTLAGGQSLKSGVTLVAKSPSQSGTLLMDAGT-----TLE 479  
Qy 571 NLHLSLSSLLANNAVTPNPPAODSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRY 630  
Db 480 TADGSLSIICSOQCRPLK-----RQEXTLKATQASQTVTLSSLSLVDPSGNVYEDVS 532  
Qy 631 WLGSNOKINVLKQLGKTKPPANAP--SDLT---LGNEMPKYGYQGSWKLAWDPTNANGPY 686  
Db 533 W--NNPOV--FSCUTLTADDPANIHITDLAADPLEKNPIHWGYQGNWALSQWEDTATKSK- 588  
Qy 687 TLKATWTKTGNPGRVAVSLVPSNLWSGILDIRSAHSAIQASVDGRSYCRGLWVGVS 746  
Db 589 AATLTWTKTGNPNPERRGTLVANTLWGSFVDVRSIQQLVATKVRSQOETRGICWCEGIN 648  
Qy 747 FFVHRRDALCGQYRISIGSYISLGANSYFGS--SMFGLAFTVEVGRSKDYVVCRSNHHACIG 805  
Db 649 FFHKDSKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKORDHPINKRASAYA 708  
Qy 806 SVYLSQQALCG---SYLFG-----DAFIRASYGFGNQHKMTSYTFAEBSDVVRD 852  
Db 709 SLHLOHLATLSPSLRYLPGSEBQVLFDAQI--SYIYSKNTMKTYTQAPKGESSW 766  
Qy 853 NNCLEBIGAGLPVITPISKLYLNLRLPFVQAEFSYADHESFTEBGDQ--ARAFKSGHLLN 911  
Db 767 NDGCALELASSLPHTALSHEGLPHAYFPFIKVEASYIHQDSPKERNITLVRSPDSGLIN 826  
Qy 912 LSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAFHLARHG 971  
Db 827 VSPVIGITFERFSRNERASYEATVIYVADVVRKNPDCCTTALLINNTSWKTGTNLRSQA 886

Qy	972	VVRGSMYASLTSNIEVYGHGRVEYRDASRGYGLSAGSRVRF	1012
		:   : : :   : : :   :   :   :   :	
Dd	887	IGRAGIFYAPSPNLEVTNSLMEIRGSSRSYNADLGKQFQF	927

```

RESULT 31
US-09-438-185A-446
Sequence 446, Application US/09438185A
Patent No. 6822071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-000411US
CURRENT APPLICATION NUMBER: US/09/438,185A
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 446
LENGTH: 1414
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
FEATURE:
OTHER INFORMATION: Cpn0444
US-09-438-185A-446

```

Query Match 18.7%; Score 985.5; DB 2; Length 1414;  
Best Local Similarity 28.0%; Pred. No. 1.4e-76;  
Matches 302; Conservative 148; Mismatches 404; Indels 226;

Qy	51	IGDP8GTTVFSAGELTTLKNLDNSI---AALPLSCFNGLLGSFTVLGRGHSITFFENIRTS	107
Db	443	VSDKTGLSIANNQEVSLT\$NAA\$TSGGAIYATKC-----TLTGNG-SLTFDQNTAGT	493
Qy	108	NGAALSNSAADGLFTIISGFKEZLSPNCNSLLAVI\$PAAT\$NKG\$QTPTTSTSPNGTIYSK	167
Db	494	SGGAIYTTETD--FTLTG-----STGTVTFTNTAKTGALYSK	530
Qy	168	TDLLLLNNEKSPFY\$NLSV\$GD-----GGAI-----DAKSLTVCGI-----SKLCVF	208
Db	531	GNN\$LSGNTNLLFSGNKATGFSN\$S\$ANQ\$CGGAIL\$FL\$S\$ASV\$TKGLMIEN\$ENVSL	590
Qy	209	Q\$NTAQADGGACQVVT\$FS\$AMANEAP\$AFVAN\$VAGRVGGI\$AAVQD\$GGQV\$SSST\$	268
Db	591	SGNTATV\$SGGA--IYATKCALHGNTLT\$FDGNTARTAGAI-----YTEIEDF	636
Qy	269	VV\$PSRNTAV\$PDGN\$VARVGGIY\$Y\$GNVAF\$NNGKTL\$FLN\$V\$APV\$YIAAKQ\$T\$SQAS	328
Db	637	TLTG\$TGT-VT\$FTNTAKTAGLHTKGN\$T\$PTKNKALV\$SGNSA-----TATAT	684
Qy	329	NT\$NNYGDGGAIFC-----KNGAQ\$G\$NNSG\$VS-----FDG\$GV	363
Db	685	TTT\$DQ\$CGGAILN\$ESDIATKSLTLT\$ENESL\$FINNTAKR\$GGGIYAPKCVISGES	744
Qy	364	V\$P\$SN\$V\$A\$G\$G\$GAIYAKCL\$SV\$NC\$GPQ\$PL\$NI\$AND\$G\$AIY\$LG\$E\$SL\$S\$AD\$GDI\$P	423
Db	745	INF\$D\$NTA\$T\$SGGAIY\$KNLSITANG\$PV\$FTN\$SGG\$KGAIY\$AD\$G\$E\$SL\$E\$AD\$GDI\$TP	804
Qy	424	DGNL\$RTAK\$EN\$AAD\$VNGTVT\$SSQAI\$SG\$SG\$GKI\$TTL\$RAK\$AGH\$QL\$FND\$PI\$EM--	478
Db	805	SGN---RAT\$GTS\$TPN-----SIHLGAGAKITKLAAP\$GHTIYFYD\$PITWEAP\$SGT	854
Qy	479	-----NNQ\$P\$Q\$SKLLK\$IND\$G\$YTG\$D\$IVF-----ANG	506
Db	855	IEELVINPV\$KAI\$VPP\$OPK\$NP\$TAS\$VP\$VP\$V\$P\$V\$P\$AN\$PNT\$GTVI\$F\$SS\$GKL\$P\$Q\$D\$ASIP\$ANT	914

Qy	507	SSTLYQNVTTIEQGRIVILREKAKLSVNSLSQTGGS-LYMEAGSXTWDFVTPPOPPQPPAANQ	565
Db	915	TTILNQINLAGNVVLKEGATLQVYSFTQOPDSTVPMDAQITLLETTTN-----NTDG	968
Qy	566	LITLSNLHLSSLSLANNAVTNPTNPQAQDSHPAVIGSTAGSVTISGPIFFEDLDDTA	625
Db	969	SIDLKNLISVNLDALDGRMIT-----IAVNSTGGLKISGLDKFHNNEGSF	1014
Qy	626	YDRYDMLGSNQKINVLKQL-GTKPPAN-----APSDLTLGNEMPKYGYGSKWLAWDP	678
Db	1015	YDN---PGLKANLPLPDLJSSGTGNLDDPNPISMSMA-----APDYGQSGMTLV--P	1065
Qy	679	NTANNGPYTLKATWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRG	738
Db	1066	KVGAGGKVTLVAEHQALGYTPKPELRATLVPSNLWAWYNIHQEIATAMSDAPSHPG	1125
Qy	739	LWVSGVSNPFYHDDALQOGYRYISGYSIGANSYFGSS-MFLGAPTEVFGRSKDYVVCVR	797
Db	1126	IWIGGIGNAFHQDKOKENAGFRLISRGYIVGGSMTTPQEVTFVAFPSQLFGSKDYVVSD	1185
Qy	798	SNHHACISGYVLSTQOALC--GSYLP-----GDAPF-----RASYPG	832
Db	1186	IKSOVYAGS-----LCAQSSYVTPHLHSSLRHVLKSVLPBLPGETVPLVHGQVSYGR	1237
Qy	833	GNQHMKTSYTFABESDVRWNNCLAGLIGAGLPVITPCKLYINELRPPFQOAFSPSYADHE	892
Db	1238	NHHNMVTKLANNTQKSDMDSDSHFAVEVGSGLSPVDL--NTRYLTSYSPYVKLQVSVNQK	1295
Qy	893	SFTEEGDQARAFKSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAVRTISGTETTL	952
Db	1296	GFQEVAADPRIFDASHLVNVVSI PMGLTFKHESAKPFSALLTLGYAVDAYDRDPHCLTS-	1354
Qy	953	LSHQETWTTTDAFLARHGHHVVVRGSMYASLTSNIENVYGHGREYRDASRGYGLSAGSRVF	1012
Db	1355	LTNCTGWSFTATNLSROAPPABASGHLKLLHGLDCFASGSCELRSSRSYNNACGTRYSE	1414

```

RESULT 32
US-09-438-185A-17
; Sequence 17, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 780
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0015
US-09-438-185A-17

```

	Query Match	18.6%	Score 980.5	DB 2	Length 780
	Best Local Similarity	29.3%	Pred. No. 1.4e-76		
	Matches	261	Conservative 142	Mismatches 320	Indels 167
	Gaps	26			
Qy	175	NEKFSFYNLVSGGGDAIKSLTVQGISKLCVPEQNTAQDGCACQVWTSFSAMANEAP	234		
Db	6	NVSLLPKSNFTSTONGGAIKATKLSLTGTTMSALPSENTSSKKGAIOTSDALTTTNGOE	65		



```
Db 718 FRYASRPNNVNGIISRRPLQNLPHFLPCLAYGHATNDKTDYANFPMVKNWNRNCWA 777
Qy 858 GBIGAGLPVITPPSKLYNLNELPFPVQAFSPYADHESFTEEGDQAPAKSGHLLNLSPVVG 917
Db 778 IECGSMPLLVFENGRLFOGAIPFMKQLVYAYQGFKEKTTADGRRFSGSLTSISVPLG 837
Qy 918 VKPDRCSSTHPKYSFMAIYICDAVRTISGTETTLSSHQETWTTDAFLARHGTVVVGSM 977
Db 838 IRPEKALSDQVLYDFSFYSIPIDFKOPSCAALVIGSDSLVPAAHVSRHAFVGSOTG 897
Qy 978 YASLTNSIEVYHGRVEYRDASRGYGLSAGSRVRF 1012
Db 898 RYHFNDEYTELLCRGSIIECRPHARNYINCGSKFRF 932

RESULT 34
US-09-198-452A-15
; Sequence 15, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 15
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; NAME/KEY: SITE
; LOCATION: 1...922
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-15

Query Match 18.2%; Score 958; DB 2; Length 922;
Best Local Similarity 27.8%; Pred. No. 1.8e-74;
Matches 293; Conservative 145; Mismatches 381; Indels 236; Gaps 32;

Qy 31 MYPQIYDGETLVSPFTVIGDPGTTVFASAGELTLKVLNLSIAALPLSCFNLGSGFT 90
Db 31 LTPEDSFHGDSONAERSYNNV---QAGDVYSLTGDVSIISNVDS--ALNKACFXVTSGSVT 85
Qy 91 VLGRGHSLTFTENIRTS--NGAALNSAADGLFTIEGPKELSPSCNCSLLAVLPAATNK 148
Db 86 FAGNHGXYFNMISSGTTKEGAVLCCQDPQATARFSGFSTLSF----- 128
Qy 149 GSQPTPTTSTPSNGTIYKTDILLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVF 208
Db 129 -NQSP--GDIEQGLCYKSNALMLNNYVVRPEQNSKTKGALSGANVTIVGVNDSVSP 185
Qy 209 QENTAAQDGCACQVVTSPFSAMANEAPIAFVANAGVGGGIAAVQDQGGVSSSTSTBDP 268
Db 186 YQNAATFGG-----AIHSGPLQIAVQAIR----- 212
Qy 269 VVPSERNTAVEPDGNVARVGGIYSGNVAFNLNGKTLPLNNVASPVVIAAKQPTSGQAS 328
Db 213 ---FAQNTAK--NGS---GGALYDGDIDIDQNAIYVLFREALUTTAI----- 252
Qy 329 NTSNNYDGGGAIFCKNGAQAGSNNSGS-----VSFDGEGVVFFSSNVAAAGGAIYAKK 382
Db 253 -----GKGAVCC-----LPTSGSSTPPVIVTFSDNKQLVPERNHSIMGGGAIYARK 299
Qy 383 LSVANGCPVQFLRNIA-----NDGAIYLGSGGELSLSDYGDIIIFDGNLKRKTAKENAAD 437
Db 300 LSISSGGPTLFINNISYANSQNLGAIADTGTGEISLSAEKGTITTFQGN-----RTSLPF 354
Qy 438 VNGVTVSSQALSMGSGGKITTLRAKAGHOILFNDPIEMANGNNQPAQSSKLLKLN---DG 494
```

```
355 LAGHILLQNA-----KFLKLOARNGYSIEFYDPI-----TSEADSGSTQININGDPKN 401
Qy 495 EGYTGDIVPA-----NGSSTLYQNVTIEQRIVLREKAKLSVNSLSQTGGS--LYM 543
Db 402 KEYTGTILFSGEKSLANDPRDFKSTIPQNVNLSAGYLVIKEGAVTVVSKPTQSPGSHLV 461
Qy 544 BAGSTWDFVTPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPPTNPPPAQDSDHPAIVG 603
Db 462 DLGT-----KLITASKEDIAITGLAIDISL-----SSSSTAATAVIK 496
Qy 604 STTAG--SVTISGPPIF-----PEDLDDTAYDRYDMLGSNQKINVLKLGITKPPANAP 654
Db 497 ANTANKQISVTDSTELISPTGNAYEDLR-----MRNSQTTPPLLSLE-----PGAGS 543
Qy 655 SDLTLGNEM-----PKYGYQGSWKLAWDPTANNNGPYTLKATWTKGTGYNPGRPRVSLVPS 711
Db 544 VVTAGDPLPVSPHYGPGQNKWLAWTGCKNVGEP-----PMDKINYKPRPEKGNLVPI 599
Qy 712 LMGSIILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLALGQGYRISGYSGLAN 771
Db 600 LMGNADVRSIMQVQETHASSLQTDRLGWLIDIGLNFHVSASEDNIRYRHSNGGYYVLSVN 659
Qy 772 SYFGSSMF-GIAPTEVFGRSKDVVVCRSNHACIGSVVLSTQQALCGSYLP-----GDA 824
Db 660 NEITPKHTSMAPSQLFSRDKDYAVSNNEYR-----MYL-----GSYLYQYTTSLGNI 707
Qy 825 FIRAS-----YFGNQNHMKTSYTFABESDVRVDNNCLA 857
Db 708 FRYASRPNNVNGIISRRFLQNLPHFLPCLAYGHATNDKTDYANFPMVKNWNRNCWA 767
Qy 858 GBIGAGLPVITPPSKLYNLNELPFPVQAFSPYADHESFTEEGDQAPAKSGHLLNLSPVVG 917
Db 768 IECGSMPLLVFENGRLFOGAIPFMKQLVYAYQGFKEKTTADGRRFSGSLTSISVPLG 827
Qy 918 VKPDRCSSTHPKYSFMAIYICDAVRTISGTETTLSSHQETWTTDAFLARHGTVVVGSM 977
Db 828 IRPEKALSDQVLYDFSFYSIPIDFKOPSCAALVIGSDSLVPAAHVSRHAFVGSOTG 897
Qy 978 YASLTNSIEVYHGRVEYRDASRGYGLSAGSRVRF 1012
Db 898 RYHFNDEYTELLCRGSIIECRPHARNYINCGSKFRF 922

RESULT 35
US-09-438-185A-15
; Sequence 15, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CFPn0013
US-09-438-185A-15

Query Match 17.8%; Score 939.5; DB 2; Length 866;
Best Local Similarity 27.0%; Pred. No. 6.7e-73;
```





QY 479 -----NNPAQSSKLLKINDGEGYTGDIVFA-----ANGSSTLYQNVYIRQG 519  
 Db 730 VPPPOKNGPIASVPVVPANPNTGTVFSSGKLPSQDASIPANTTTILNOKINLAGG 789  
 QY 520 RIVLREKAKLSVNSLSQTS-LYMEAGSTWDFVTPPOPPPPAANQILITLSNHLSS 578  
 Db 790 NVVLAKGATLQVYSTQPDSTVFDAGTTLTETTTN-----NTDGSIDLKNSLVNDA 843  
 QY 579 LLANNAVTPNPAPQDSHPAVIGSTAGSVTISGPIPPEDDLDTAYDRYDMLGSNOKI 638  
 Db 844 LDGKEMIT-----IAVNSTSGGLKISGDLKPHNNEGSFYDN---PGLKANL 886  
 QY 639 NVLKLQ-LGTRPPAN-----APSDTLGNEMPKYGYGSKWKLAWDPNTANNPPTLKAT 691  
 Db 887 NLPLDLSTSGTVNLDNFPNPSMA-----APDYGYGQSMTLV--PKVGAGGKVTLVAE 940  
 QY 692 WTKTGNPGPERSVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFFYHD 751  
 Db 941 WQALGYTPKPELRATLVNLSLWNAVNIHSIQEITATMSDAPSHPGIWIIGIGNAFQD 1000  
 QY 752 RDALGQGYRISGGYSLGANSYFGSS-MFLAFTVFGRSKDYVVCRSNHHACIGSVYLS 810  
 Db 1001 KQKENAGFLLISRGYIVGSGMTTPQEXTFAVAFSOLFQSKDYVVSIDIKSOVYAGS---- 1056  
 QY 811 TQALC--GSYLP-----GDAFI-----RASVGFQNGHMTSYTFAE 845  
 Db 1057 ----LCAQSSYVILPHSSLRRLHVLKVLPELPGTPLVHQVSYGRNHNMTTKLANNT 1112  
 QY 846 ESDVRWDNN 854  
 Db 1113 QKQSDWDH 1121

## RESULT 37

US-09-438-185A-451  
 ; Sequence 451, Application US/09438185A  
 ; Patent No. 6822071  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stephens, Richard  
 ; APPLICANT: Mitchell, Wayne  
 ; APPLICANT: Kalman, Sue  
 ; APPLICANT: Davis, Ronald  
 ; APPLICANT: The Regents of the University of California  
 ; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence  
 ; FILE REFERENCE: 018941-00041105  
 ; CURRENT APPLICATION NUMBER: US/09/438,185A  
 ; CURRENT FILING DATE: 2002-03-13  
 ; PRIOR APPLICATION NUMBER: US 60/108,279  
 ; PRIOR FILING DATE: 1998-11-12  
 ; PRIOR APPLICATION NUMBER: US 60/128,606  
 ; PRIOR FILING DATE: 1999-04-08  
 ; NUMBER OF SEQ ID NOS: 1074  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 451  
 ; LENGTH: 634  
 ; TYPE: PRF  
 ; ORGANISM: Chlamydia pneumoniae  
 ; FEATURE:  
 ; OTHER INFORMATION: Cpn0449  
 US-09-438-185A-451

Query Match 15.3%; Score 808; DB 2; Length 634;  
 Best Local Similarity 30.8%; Pred. No. 1.3e-61;  
 Matches 215; Conservative 107; Mismatches 265; Indels 110; Gaps 19;

QY 358 FDGEVVPFSSNVAAGKGAIVAKKLSVANCQPVQFLNIANDGGAIVLGESELSLSAD 417  
 Db 6 FRGGGGSFNNIVQ-----TTAGNGGAISILAAEGCSLSAE 43  
 QY 418 YGDIIFDGN-----LKRITAKENADVNGVTSSQAIMSGSGKITTLLRAKAGHQILFNDDP 472  
 Db 44 AGDITFNGNAIVATTPQTTKNSIDI-----GSTAKITNLRAISGHSIFFFYDP 91

QY 473 IEMANGNNQPAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNV 514  
 Db 92 ITA-----NTAADSTDTLNLAKADAGNSTDYSGSVFSGEKLSEDEAKVADNLJTSLKQPV 147  
 QY 515 TIEQGRIVLREKAKLSVNSLSQTS-GLSLYMEAGSTWDFVTPPOPPPPAANQILITLSNHL 573  
 Db 148 TLTAGNLVLKRGVTLDTKGTQTAGSSVIMDAGTTL-----KASTEETVLTLS 196  
 QY 574 LSLSLANNAVTPNPAPQDSHPAVIGSTAGSVTISGPIPPEDDLDTAYDRYDMLG 633  
 Db 197 IPVDSL-----GEGKKVVIASAASKVALSGPILLDDNOGNAYENHD-LG 241  
 QY 634 SNQKINVLKQ-LGTRPPANAPSDTLGNEMPKYGYGSKWKLAWDPNTANNPPTLKAT- 691  
 Db 242 KTODFSFVQLSALGTATTTDVPAPVTVATP--THYGYGQGTWGMTWVDVDTAST-PKTKTATL 299  
 QY 692 -WTKTGNPGPERSVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFFYH 750  
 Db 300 ANWTGVLNPNRQCPPLVPNSLWGSFSDIQAIQGVIERALSALTLCSDRGFMAAGVANFLDK 359  
 QY 751 DRDALGQGYRISGGYSL-GANSYFGSSMFLAFTVFGRSKDYVVCRSNHHACIGSVYL 809  
 Db 360 DKGEKKRYRHSKGYAIGGAATCSENLSIFAFQOLFQSKDYVVAKNHTDTYAGAFYI 419  
 QY 810 S-----TQALCGSYLFGDAFI--RASVGFQNGHMTSYTFAESDVRWDNNCL 856  
 Db 420 QHITCSEGFICGLADKLPGSMHKLPLVLEQLAYSHVSNDLTKYTYAYPEYKVGWGNNAF 479  
 QY 857 AGEIGAGLPIVITPSKLY-LNELRPPFVQARFSAVDHESPTTEGQARAPKSGHLLNLSVP 915  
 Db 480 NMLGASSHSY--PEYLHCFDTYAPYIKMLTYIRQDSFSEKTEGSRFDDSNLFLSLP 537  
 QY 916 VGVKFDRCSTHPNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAFHLARHGTVVRG 975  
 Db 538 IGKVFKEFSDCNDPSYDLTSLVYVPLIRNDPKCTTALVISGASWETVANNLARQALQVRA 597  
 QY 976 SMYASLTSNIEVYCHGRYEYRDASRGYLSAGSRVP 1012  
 Db 598 GSHYAPSPMEVLGQPVFEVRGSSRIYNVDLGGKQF 634

## RESULT 38

US-09-198-452A-474  
 ; Sequence 474, Application US/09198452A  
 ; Patent No. 6559294  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Grifflais, R.  
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
 ; TITLE OF INVENTION: and treatment of infection  
 ; FILE REFERENCE: 9710-003-999  
 ; CURRENT APPLICATION NUMBER: US/09/198,452A  
 ; CURRENT FILING DATE: 1998-11-24  
 ; NUMBER OF SEQ ID NOS: 6849  
 ; SEQ ID NO 474  
 ; LENGTH: 643  
 ; TYPE: PRF  
 ; ORGANISM: Chlamydia pneumoniae  
 US-09-198-452A-474

Query Match 15.3%; Score 804; DB 2; Length 643;  
 Best Local Similarity 30.8%; Pred. No. 3.1e-61;  
 Matches 214; Conservative 107; Mismatches 264; Indels 110; Gaps 19;

QY 360 GEGVVPFSSNVAAGKGAIVAKKLSVANCQPVQFLNIANDGGAIVLGESELSLSADYG 419  
 Db 17 GGGGGSFNNIVQ-----TTAGNGGAISILAAEGCSLSAEAG 54  
 QY 420 DIIFDGN-----LKRITAKENADVNGVTSSQAIMSGSGKITTLLRAKAGHQILFNDDP 474  
 Db 55 DITFNGNAIVATTPQTTKNSIDI-----GSTAKITNLRAISGHSIFFFYDPIT 102  
 QY 475 MANGNNQPAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNV 516

Db 103 A-----NTAADSOTDNLKADAGNSTDYSGSVFSEKLSSEDAKVADNLTSLAQPVTL 158  
 QY 517 EGRIVLRKAKLSVNSLSQT--GGSLYMEAGSTWDFVTPPPQPPQPPAANQLITLNLHLS 575  
 Db 159 TAGNLVLRKGVTLDTKGTQTAGSSVIMDAGTTL-----KASTEETVLTGLSIP 207  
 QY 576 LSSLLANNATVPPTNPQADSHPAVIGSTTAGSVTISGPIPFEDLDDPAYDRYDNLGSN 635  
 Db 208 VDLSL-----GEGKKVJAASAARKNVALSGPILLLDQGNAYENHD--LGKT 252  
 QY 636 QKINVLKQ--LGTKPPANAPSOLDLTGNEMPKYQOGSKWLAMPNTANNPPTLTKAT--W 692  
 Db 253 QDPSFVLSALGTATTTDVPVATP--THGYQGTWGTWDDTFAST--PKTKATLAW 310  
 QY 693 TKTYNPGPERSVASLPNSLWMSILDIRSAHSAIQASVDGRSVCRLWVGVSNSPFYHDR 752  
 Db 311 TWTGYPNPERQPLVPNSLWMSFSDIAIQGVIERALSALTCLSDRGFAAGVANFLDKDK 370  
 QY 753 DALGQGYRVISGYSL--GANSYFGSSMFGLAFTVFRGSKDYVVCRSNHHACIGSVYLS- 810  
 Db 371 KGEKRYRKHSGGYATGGAAQTSENLSIFAPCOLFGSKDPLVAKNHTDTTAGAFYIOH 430  
 QY 811 -----TQALCGSYLFGDAFI--RASVGFQGNHMKTSYTFABESDVRMNNCLAG 858  
 Db 431 ITCSCFIGCLLDKLPKGSWSHKLVLLEGQLAYSHVNDLTKTYTAYPEVKGSGWNAFNM 490  
 QY 859 EIGAGLPPIVITPSKLY--LNELRPFVQAFPSYADHESFTREGDQARAFKSHLLNLNVPG 917  
 Db 491 MLGASSHSY--PEYLHCFDTPYAVIKLNTYIRQDSFSEKTEGRSFDSDNLFNLSLPIG 548  
 QY 918 VKPDRCSSTHPNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGVVVRGSM 977  
 Db 549 VKPEKFSDCNDPSYDLTSLVYPLIRNDPKCTALTALVISGASWETYNANLARQALQVRAGS 608  
 QY 978 YASLTSNIEVYGHGREYRSDASRGYGLSAGSRVRF 1012  
 Db 609 HYAFSPMFEVLQGFVFEVRGSSRIYVNDLGKQF 643

RESULT 39

US-09-556-877-175  
 ; Sequence 175, Application US/09556877  
 ; Patent No. 6432916  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Probst, Peter  
 ; APPLICANT: Bhatia, Ajay  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Fling, Steve  
 ; APPLICANT: Maisonneuve, Jeff  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
 ; FILE REFERENCE: 210121.469C5  
 ; CURRENT APPLICATION NUMBER: US/09/556,877  
 ; CURRENT FILING DATE: 2000-04-19  
 ; NUMBER OF SEQ ID NOS: 305  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 175  
 ; LENGTH: 880  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1)...(880)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-556-877-175

Query Match 14.6%; Score 769.5; DB 2; Length 880;  
 Best Local Similarity 26.2%; Pred. No. 5.5e-58;  
 Matches 264; Conservative 146; Mismatches 395; Indels 203; Gaps 33;

QY 53 DPSGTTVPFSAAGELTKNLDNSIALPLSCFNLL-----GSFTVLGRGHSITFE 101

Db 28 DPLGETA-----LTKPNPNHVCTFFEDCTNWESLFPALCAHASQDDPLYLVGNSYCWFS 82  
 QY 102 NIRTSTNGAALSAAADGLFTIEGKELSFNCNCSLLAVLPAATTNKGSQPTTTSTPSN 161  
 Db 83 KLHITDPKEALFKKGD--LSIQNFRFLSFDDCS-----KSSPS- 121  
 QY 162 GTIYSKT--DLLLNNKPFYSNLSVSGGGDAIDAKSLTVQGISKLCVQENTAOADGAC 220  
 Db 122 -IIHQKQGLSLRNNGSMFPCRNHAEKSGGSAISADAFSLQHNLYLFTAFSENSKNGGAI 180  
 QY 221 QVITSFSAWANEAPAFVANAGVGGGIAAVQDQCGQSVSSSTSTEDPVVFSRNTAVEF 280  
 Db 181 QAQT--FSLSRVSPISFARNRADLNGGAICC----- 210  
 QY 281 DGNVARVGGGYSYGVNVAFLNNGKTLFLNNVASPVYIAAKQPTSCQASNTSNYDGGAI 340  
 Db 211 -----SNLCSGNV-----NPLFF-----TGNSATNGGAI 235  
 QY 341 FCKNGAAGSNNSVSPDGEVVPFSSNVAAGKGGAIYAKKLSVANGCPVQFLNIAND 400  
 Db 236 CCI--SDLNTSEKGLSLACNQETLTFASNAKEKGGAIYAKHMLVRYNGPVSFINNSAKI 293  
 QY 401 GGAIVLGESGELSADYGDIIFDGNLAKRTAKENADVNGVTVSQAISMGSGGKITTLR 460  
 Db 294 GGAIAIQSGGSLILAGEGSLFQNNQSRTSQGLVR--NAIYLXKDAI-----LSSLE 345  
 QY 461 AKAGHQIILFNDPI--EMANGNNOAQSLLKLKINDGEGVTG-----DIVFA----- 504  
 Db 346 ARNG--DILFPDPIVGESSKESPLSSLOASTSTPTATASPLVIQTSANRSVIFSSERL 404  
 QY 505 -----NGSSTLYQNVITIEQGRIVLRKAKLSVNSLSQGTGS--LYMEAGSTWDFVTPQ 555  
 Db 405 SEEEKTPDMLTSOLQOPIELKSGRLVKDRAVLSAPLSQDPOALLIMEAGTS----- 457  
 QY 556 PQQPPAANQLITLNLHLSLSSLANNAVTPPTNPQADSHPAVIGSTTAGSVTISGP 615  
 Db 458 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTIHAP 487  
 QY 616 -----IPEDL--DDTAYDRYDNLGSKNOKINVLKQLQTKPPANAPSDITLGNEMPKY 668  
 Db 488 NLSIQKIFLUSGSDENFYENVLLSKEQN-----NPLLTLPKEQSHLHLPDGNLSSHFGY 543  
 QY 669 QGSKWLAMPNTANNPPTLTKATWTKTYNPGPERVASLPNSLWMSILDIRSAHSAIOA 728  
 Db 544 QGDWTFSM--KDSDEG--HSLIANWTPKNVPHPERQSTLVANTLWNTYSDMQAVQSMINT 600  
 QY 729 SVDRGSYCRGLWVSGVSNFFY--HDDRDLG--QGYRISGGYSGIGANSY--FGSSMFGLAFT 785  
 Db 601 TAHGGAYLFGTWGSAVSNLFYVHDSGKPIDNWHHRSGLYFGISTHSLDDHSHFCLAAQ 660  
 QY 786 VFGRSKDYVVCRSNHHACIGSVYLTQOALCGSYLFGDAPIRASVYFGNQHMKTSY--TFA 844  
 Db 661 LIGKSSDSFITSTE-----TTSYIATVQALATSLM--KLSAQACYNESHELTKYRSP 714  
 QY 845 EESDVRMNNCLAGEITGAGLPVITPESKLYLNELRPFVQAFPSYADHESFTREGDQARAF 904  
 Db 715 KEGFGSWHSAVSGEVCASIPVNSGSLF--SSFSIFSKLQGFSGTQDGFSESSGIRSF 773  
 QY 905 KSGHLLNLVSVGVKPFDRCSSTHPNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAF 964  
 Db 774 SASSFRNISLPIGITEPKKSQKTRTYTYFLGAYIQDLKRDVSGSPVLLKNAVSWDAPMA 833  
 QY 965 HLABHGVVVRGSMYASLTSNIEVYGHGREYRSDASRGYGLSAGSRVRF 1012  
 Db 834 NLDSRAYMFLTNQRAL--HRLQTLNLNVSVCLRGQSHSYSLDUGTTTTRF 880

RESULT 40

US-09-620-412C-175  
 ; Sequence 175, Application US/09620412C  
 ; Patent No. 6448234  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Steven P. Fling



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:31:04 ; Search time 168 Seconds  
(without alignments)  
2516.924 Million cell updates/sec

Title: US-10-701-844-2  
Perfect score: 5267  
Sequence: 1 MQTSFHKFLSMILAYSCS.....YEURDARGYGLSAGSRVRF 1012

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	5267	100.0	1012	4	US-10-701-844-2
2	5267	100.0	1012	4	US-10-766-711-2
3	5267	100.0	1012	5	US-10-931-779-2
4	5131.5	97.4	1013	4	US-10-467-534-9
5	5131.5	97.4	1013	5	US-10-498-327-5
6	5131.5	97.4	1013	5	US-10-498-327-81
7	5125.5	97.3	1013	5	US-10-498-327-83
8	5123.5	97.3	1013	5	US-10-498-327-9
9	5123.5	97.3	1013	5	US-10-498-327-13
10	5123.5	97.3	1013	5	US-10-498-327-93
11	5121.5	97.2	1013	5	US-10-498-327-1
12	5121.5	97.2	1013	5	US-10-498-327-3
13	5116	97.1	984	4	US-10-701-844-43
14	5116	97.1	984	4	US-10-766-711-43
15	5116	97.1	984	5	US-10-931-779-43
16	5112.5	97.1	1013	5	US-10-498-327-17
17	5106.5	97.0	1013	4	US-10-701-844-15
18	5106.5	97.0	1013	4	US-10-701-844-16
19	5106.5	97.0	1013	4	US-10-766-711-15
20	5106.5	97.0	1013	4	US-10-766-711-16
21	5106.5	97.0	1013	5	US-10-931-779-15
22	5106.5	97.0	1013	5	US-10-931-779-16
23	5090	96.6	1006	3	US-08-841-132-190
24	5090	96.6	1006	5	US-10-872-155-190
25	5084	96.5	982	3	US-09-841-132-176
26	5084	96.5	982	5	US-10-872-155-176
27	3336.5	63.3	670	5	US-10-197-220-169

28	3336.5	63.3	670	6	US-11-109-468-169
29	2552	48.5	505	4	US-10-701-844-17
30	2552	48.5	505	4	US-10-766-711-17
31	2552	48.5	505	5	US-10-931-779-17
32	2350	44.6	458	4	US-10-701-844-36
33	2350	44.6	458	4	US-10-766-711-36
34	2350	44.6	458	5	US-10-931-779-36
35	1735	32.9	325	4	US-10-701-844-37
36	1735	32.9	325	4	US-10-766-711-37
37	1735	32.9	325	5	US-10-931-779-37
38	1365.5	25.9	973	4	US-10-312-273-3
39	1359.5	25.8	973	4	US-10-352-618-2
40	1128	21.4	928	3	US-09-428-122-2
41	1122	21.3	926	3	US-09-738-269-57
42	1122	21.3	926	4	US-10-023-437-57
43	1091	20.7	928	4	US-10-312-273-95
44	1090	20.7	928	4	US-10-312-273-115
45	1090	20.7	928	4	US-10-282-122A-54681

ALIGNMENTS

RESULT 1

US-10-701-844-2  
; Sequence 2, Application US/10701844  
; Publication No. US20040067524A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/10/701,844  
; CURRENT FILING DATE: 2003-11-04  
; PRIOR APPLICATION NUMBER: US/09/612,402B  
; PRIOR FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 1012  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-701-844-2

Query Match	100.0%;	Score 5267;	DB 4;	Length 1012;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1012;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MQTSFHKFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPPTVIGDPSGTTVF	60	
Db	1	MQTSFHKFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPPTVIGDPSGTTVF	60	
QY	61	SAGELTKMLDINSIAALPLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSNSAADGL	120	
Db	61	SAGELTKMLDINSIAALPLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSNSAADGL	120	
QY	121	FTIEGPKELSPFNCNLSLLAVLPAATNKGSGTPTTTSTPSNGHIYSKTDLLLNNEKFSF	180	
Db	121	FTIEGPKELSPFNCNLSLLAVLPAATNKGSGTPTTTSTPSNGHIYSKTDLLLNNEKFSF	180	
QY	181	YSLNLSGSGDAIDAKSLTVQGISKLCVQENTAQDGGACQVVTFSAMANEAPIAFVAN	240	
Db	181	YSLNLSGSGDAIDAKSLTVQGISKLCVQENTAQDGGACQVVTFSAMANEAPIAFVAN	240	
QY	241	VAGVRGGIAAVODGQGVSSSTSTEDPVVPSRNTAVEFDGNVARVGGGIYSYGNVAF	300	
Db	241	VAGVRGGIAAVODGQGVSSSTSTEDPVVPSRNTAVEFDGNVARVGGGIYSYGNVAF	300	
QY	301	NNGKTLFLNNVSPVYIAAKQPTSGQASNTNNYDGGGAI FCKNGAQAGSNNSGSVSFDG	360	
Db	301	NNGKTLFLNNVSPVYIAAKQPTSGQASNTNNYDGGGAI FCKNGAQAGSNNSGSVSFDG	360	

QY	361	EGVVPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSGBLSLSADYGD	420
DB	361	EGVVPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSGBLSLSADYGD	420
QY	421	IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN	480
DB	421	IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN	480
QY	481	QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS	540
DB	481	QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS	540
QY	541	LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA	600
DB	541	LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA	600
QY	601	VIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNOKINVLKQLGTGKPPANAPSDLTG	660
DB	601	VIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNOKINVLKQLGTGKPPANAPSDLTG	660
QY	661	NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLWGSILDIR	720
DB	661	NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLWGSILDIR	720
QY	721	SAHSAIQASVDGRSYCRGLMWGVSNSFFYHRRDALGQGYRIVISGYSLGANSYFGSSMFG	780
DB	721	SAHSAIQASVDGRSYCRGLMWGVSNSFFYHRRDALGQGYRIVISGYSLGANSYFGSSMFG	780
QY	781	LAPTEVFRGSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA	840
DB	781	LAPTEVFRGSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA	840
QY	841	YTFABESDVRWNNCLAGEIGAGLPIVITPSKLYLNELRPFVQAEFVSADHESFTEEGDQ	900
DB	841	YTFABESDVRWNNCLAGEIGAGLPIVITPSKLYLNELRPFVQAEFVSADHESFTEEGDQ	900
QY	901	ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT	960
DB	901	ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT	960
QY	961	TDAPFLARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF	1012
DB	961	TDAPFLARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF	1012
RESULT 2			
US-10-766-711-2			
; Sequence 2, Application US/10766711			
; Publication No. US20040137005A1			
; GENERAL INFORMATION:			
; APPLICANT: Jackson, W. James			
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof			
; FILE REFERENCE: 7969-062-999			
; CURRENT APPLICATION NUMBER: US/10/766,711			
; PRIOR FILING DATE: 2004-01-27			
; PRIOR APPLICATION NUMBER: US/08/942,596			
; PRIOR FILING DATE: 1997-10-02			
; NUMBER OF SEQ ID NOS: 43			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 2			
; LENGTH: 1012			
; TYPE: PRT			
; ORGANISM: Chlamydia sp.			
US-10-766-711-2			
Query Match 100.0%; Score 5267; DB 4; Length 1012;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MQTSFHKEFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTVTF	60

DB	1	MQTSFHKEFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTVTF	60
QY	61	SAGELTLKLNLDNSIAALPLSCFGLNLGSLFTVLGRGHSITFTENIRTSNGAALSNSAADGL	120
DB	61	SAGELTLKLNLDNSIAALPLSCFGLNLGSLFTVLGRGHSITFTENIRTSNGAALSNSAADGL	120
QY	121	FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTSTPNSGTIYKTDLLLLNNKPSF	180
DB	121	FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTSTPNSGTIYKTDLLLLNNKPSF	180
QY	181	YSNLVSGDGAIDAKSLTVQGISKLCVPOENTAOADGACQVVTSPSAMANEAPIAVAN	240
DB	181	YSNLVSGDGAIDAKSLTVQGISKLCVPOENTAOADGACQVVTSPSAMANEAPIAVAN	240
QY	241	VAGVGGGIAAQQGQGGSSSTSTEDPVVSPFSRNTAVEFDGNVARYGGGIYSYGNVAF	300
DB	241	VAGVGGGIAAQQGQGGSSSTSTEDPVVSPFSRNTAVEFDGNVARYGGGIYSYGNVAF	300
QY	301	NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGDGGAIFCXGGAQAGSNNSGVSFDG	360
DB	301	NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGDGGAIFCXGGAQAGSNNSGVSFDG	360
QY	361	EGVVPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSGBLSLSADYGD	420
DB	361	EGVVPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSGBLSLSADYGD	420
QY	421	IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN	480
DB	421	IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN	480
QY	481	QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS	540
DB	481	QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS	540
QY	541	LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA	600
DB	541	LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA	600
QY	601	VIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNOKINVLKQLGTGKPPANAPSDLTG	660
DB	601	VIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNOKINVLKQLGTGKPPANAPSDLTG	660
QY	661	NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLWGSILDIR	720
DB	661	NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLWGSILDIR	720
QY	721	SAHSAIQASVDGRSYCRGLMWGVSNSFFYHRRDALGQGYRIVISGYSLGANSYFGSSMFG	780
DB	721	SAHSAIQASVDGRSYCRGLMWGVSNSFFYHRRDALGQGYRIVISGYSLGANSYFGSSMFG	780
QY	781	LAPTEVFRGSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA	840
DB	781	LAPTEVFRGSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA	840
QY	841	YTFABESDVRWNNCLAGEIGAGLPIVITPSKLYLNELRPFVQAEFVSADHESFTEEGDQ	900
DB	841	YTFABESDVRWNNCLAGEIGAGLPIVITPSKLYLNELRPFVQAEFVSADHESFTEEGDQ	900
QY	901	ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT	960
DB	901	ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT	960
QY	961	TDAPFLARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF	1012
DB	961	TDAPFLARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF	1012
RESULT 3			
US-10-931-779-2			
; Sequence 2, Application US/10931779			
; Publication No. US20050048557A1			
; GENERAL INFORMATION:			
; APPLICANT: Jackson, W. James			

```

; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-2

Query Match      100.0%; Score 5267; DB 5; Length 1012;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTSPHKPFLSMILAYSCCLNGGGYAAIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCCLNGGGYAAIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60

QY 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNGAALSNSAADGL 120
DB 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNGAALSNSAADGL 120

QY 121 FTIEGPKELSFNCNLSLLAVLPAATTNKGSTPTTTSTPSNGTIYKTDLLLNNEKFSF 180
DB 121 FTIEGPKELSFNCNLSLLAVLPAATTNKGSTPTTTSTPSNGTIYKTDLLLNNEKFSF 180

QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAAQDGGACQVTVSFSAMANEAPIAFVAN 240
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAAQDGGACQVTVSFSAMANEAPIAFVAN 240

QY 241 VAGVRGGIAAODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVRGGIYSGNVAF 300
DB 241 VAGVRGGIAAODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVRGGIYSGNVAF 300

QY 301 NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNNGYDGGAIKXGAGQAGSNNSGVSFDG 360
DB 301 NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNNGYDGGAIKXGAGQAGSNNSGVSFDG 360

QY 361 EGVVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 420
DB 361 EGVVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 420

QY 421 IIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLRAKAGHQLFPNDPIEMANGNN 480
DB 421 IIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLRAKAGHQLFPNDPIEMANGNN 480

QY 481 QPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGS 540
DB 481 QPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGS 540

QY 541 LYMEAGSTWDPVTPQPPQPPAANQLITLNLHLSSLANNANVTTPNPAPQDSHPA 600
DB 541 LYMEAGSTWDPVTPQPPQPPAANQLITLNLHLSSLANNANVTTPNPAPQDSHPA 600

QY 601 VIGSTTAGSVTISGPIEPEDLDTDYDRYDWLGSNQNINVLKQLGKTPPANAPSDLTG 660
DB 601 VIGSTTAGSVTISGPIEPEDLDTDYDRYDWLGSNQNINVLKQLGKTPPANAPSDLTG 660

QY 661 NEMPKYQGSWKLAWDPNTANNPYTLKATWTGTGNPGERVASLVPNSLWMSIILDIR 720
DB 661 NEMPKYQGSWKLAWDPNTANNPYTLKATWTGTGNPGERVASLVPNSLWMSIILDIR 720

QY 721 SAHSAIQASVDGRSICRGLWVSGVSNFFYHDDRDLALGQGYRIYSGYSLGANSYFGSSMFG 780
DB 721 SAHSAIQASVDGRSICRGLWVSGVSNFFYHDDRDLALGQGYRIYSGYSLGANSYFGSSMFG 780

QY 781 LAPTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIRASYGFGNOHMKTS 840

```

```

DB 781 LAPTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIRASYGFGNOHMKTS 840
QY 841 YTFAESDVRWDDNCLAGEIGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTEEGDQ 900
DB 841 YTFAESDVRWDDNCLAGEIGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTEEGDQ 900
QY 901 ARAPKSGHLLNLSVEPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960
DB 901 ARAPKSGHLLNLSVEPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960
QY 961 TDAPFLARHGTVVRGSMYASLTSNIEVYGHGREYERDASRGYGLSAGSRVRF 1012
DB 961 TDAPFLARHGTVVRGSMYASLTSNIEVYGHGREYERDASRGYGLSAGSRVRF 1012

```

RESULT 4

```

US-10-467-534-9
; Sequence 9, Application US/10467534
; Publication No. US20040131625A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verilant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/10/467,534
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-467-534-9

```

Query Match 97.4%; Score 5131.5; DB 4; Length 1013;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

```

QY 1 MQTSPHKPFLSMILAYSCCLNGGGYAAIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCCLNGGGYAAIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60

QY 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNGAALSNSAADGL 120
DB 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNGAALSNSAADGL 120

QY 121 FTIEGPKELSFNCNLSLLAVLPAATTNKGSTPTTTSTPSNGTIYKTDLLLNNEKFSF 180
DB 121 FTIEGPKELSFNCNLSLLAVLPAATTNKGSTPTTTSTPSNGTIYKTDLLLNNEKFSF 180

QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAAQDGGACQVTVSFSAMANEAPIAFVAN 240
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAAQDGGACQVTVSFSAMANEAPIAFVAN 240

QY 241 VAGVRGGIAAODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVRGGIYSGNVAF 300
DB 241 VAGVRGGIAAODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVRGGIYSGNVAF 300

QY 301 NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNNGYDGGAIKXGAGQAGSNNSGVSFD 359
DB 301 NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNNGYDGGAIKXGAGQAGSNNSGVSFD 359

QY 360 GEGVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 419
DB 360 GEGVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 419

QY 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLRAKAGHQLFPNDPIEMANGNN 479

```

```
Db 421 DIIIPGNLKRKTAKENAADVNGVTVSSQAIISMGSGGKITTIRAKAGHQILFNDPIEMANGN 480
Qy 480 NOPAQS8KLLKINDGEGYTGDIIVFANGSSTLYQNVTTI8QGRIVLREKAKLSVNSLSQTGG 539
Db 481 NOPAQS8BFLKINDGEGYTGDIIVFANGSSTLYQNVTTI8QGRIVLREKAKLSVNSLSQTGG 540
Qy 540 SLYMEAGSTWDPVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPPAQDSHP 599
Db 541 SLYMEAGSTLDPVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNQKINVLKQLGCTKPPANAPSDLT 659
Db 601 AIIIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNQKIDVLKQLGCTQPSANAPSDLT 660
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNPPTLKAATWTKTGYNPQPERVASLVPNSLMGSI 719
Db 661 GNEMPKYGYQGSWKLAWDPNTANNPPTLKAATWTKTGYNPQPERVASLVPNSLMGSI 720
Qy 720 RSAHSAIQASVDGRSVCYCRGLWVGVSNNFFYHDDRDLALGQGYRISGYSIGLANSYFGSSMP 779
Db 721 RSAHSAIQASVDGRSVCYCRGLWVGVSNNFFYHDDRDLALGQGYRISGYSIGLANSYFGSSMP 780
Qy 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIRASYGFGNQHMKT 839
Db 781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIRASYGFGNQHMKT 840
Qy 840 SYTFAESDVRWNNCLAGEIGAGLPVITPSKLYLNLRLPFPVQAFPSYADHESFTREGD 899
Db 841 SYTFAESDVRWNNCLAGEIGAGLPVITPSKLYLNLRLPFPVQAFPSYADHESFTREGD 900
Qy 900 QARAPKSGHLNLSVPVGVKFDRCSTHFNKYSFMAAYICDAYRTISGTETTLISHQETW 959
Db 901 QARAPKSGHLNLSVPVGVKFDRCSTHFNKYSFMAAYICDAYRTISGTETTLISHQETW 960
Qy 960 TTDAPHLARHGVIVRGSMYASLTSNIEVYGHGREYRDASRGYGLSAGSVRF 1012
Db 961 TTDAPHLARHGVIVRGSMYASLTSNIEVYGHGREYRDTSRGYGLSAGSVRF 1013
```

Query Match 97.4%; Score 5131.5; DB 5; Length 1013;

Best Local Similarity 97.0%; Pred. No. 0;  
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

```
Qy 1 MOTSFHKFELSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP 60
Db 1 MOTSFHKFELSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP 60
Qy 61 SAGELTLKXLDNSIAALPLSCFNLG8FTVLGRHSLTFENIRTSNTGAALSNSAADGL 120
Db 61 SAGELTLKXLDNSIAALPLSCFNLG8FTVLGRHSLTFENIRTSNTGAALSNSANSGL 120
```

RESULT 6

US-10-498-327-81

; Sequence 81, Application US/10498327

; Publication No. US20050106162A1

; GENERAL INFORMATION:

; APPLICANT: Grandi, Guido

; APPLICANT: Ratti, Giulio

; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis

; FILE REFERENCE: 002441.00085

; CURRENT APPLICATION NUMBER: US/10/498,327

; CURRENT FILING DATE: 2004-06-10

; PRIOR APPLICATION NUMBER: PCT/IB02/05761



```
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 1013
; TYPE: PR
; ORGANISM: Chlamydia trachomatis
US-10-498-327-81

Query Match          97.4%; Score 5131.5; DB 5; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 MQTSPHKPFLSMILAYSCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60

QY 61 SAGELTKNLNDSIAALPLSCFNLGSLFTVLGRGHS�FTENIRTSNGAALSNSAAGL 120
DB 61 SAGELTKNLNDSIAALPLSCFNLGSLFTVLGRGHS�FTENIRTSNGAALSNSAAGL 120

QY 121 FTIEGPKELSFNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180
DB 121 FTIEGPKELSFNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180

QY 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTVSFSANEAPIAFVAN 240
DB 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTVSFSANEAPIAFVAN 240

QY 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300
DB 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300

QY 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359

QY 360 GEGVWFFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419
DB 360 GEGVWFFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419

QY 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLIFNDPIEMANGN 479
DB 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLIFNDPIEMANGN 479

QY 480 NQPAQSSEPLKINDGEGYTGDI VFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539
DB 480 NQPAQSSEPLKINDGEGYTGDI VFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539

QY 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 599
DB 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 599

QY 541 SLYMEAGSTLDFVTTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 600
DB 541 SLYMEAGSTLDFVTTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 600

; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 1013
; TYPE: PR
; ORGANISM: Chlamydia trachomatis
US-10-498-327-83

Query Match          97.3%; Score 5125.5; DB 5; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 21; Mismatches 11; Indels 1; Gaps 1;

QY 1 MQTSPHKPFLSMILAYSCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60

QY 61 SAGELTKNLNDSIAALPLSCFNLGSLFTVLGRGHS�FTENIRTSNGAALSNSAAGL 120
DB 61 SAGELTKNLNDSIAALPLSCFNLGSLFTVLGRGHS�FTENIRTSNGAALSNSAAGL 120

QY 121 FTIEGPKELSFNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180
DB 121 FTIEGPKELSFNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180

QY 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTVSFSANEAPIAFVAN 240
DB 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTVSFSANEAPIAFVAN 240

QY 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300
DB 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300

QY 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359

QY 360 GEGVWFFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419
DB 360 GEGVWFFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419

QY 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLIFNDPIEMANGN 479
DB 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLIFNDPIEMANGN 479

QY 480 NQPAQSSEPLKINDGEGYTGDI VFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539
DB 480 NQPAQSSEPLKINDGEGYTGDI VFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539

QY 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 599
DB 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 599

QY 541 SLYMEAGSTLDFVTTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 600
DB 541 SLYMEAGSTLDFVTTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSDHP 600
```



Best Local Similarity 96.6%; Pred. No. 0;  
Matches 979; Conservative 22; Mismatches 11; Indels 1; Gaps 1;

QY 1 MQTSPHKFFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPFVTVIGDPSGTTVP 60  
DB 1 MQTSPHKFFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPFVTVIGDPSGTTVP 60

QY 61 SAGELTTLKNLONSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGAALSNSAADGL 120  
DB 61 SAGELTTLKNLONSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGAALSNSAADGL 120

QY 121 FTIEGFKELSPNCNLSLAVLPAATNKGSOPTTTTSPSNGTIYKTDLLILNNEKFSF 180  
DB 121 FTIEGFKELSPNCNLSLAVLPAATNKGSOPTTTTSPSNGTIYKTDLLILNNEKFSF 180

QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQDGGACQVVTSPSAMANEAPAFVAN 240  
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQDGGACQVVTSPSAMANEAPAFVAN 240

QY 241 VAGVGGGIAAVQDGGQGVSSSTSTEDPVVFSRNTAVFPGNVARVGGGIYSGNVAF 300  
DB 241 VAGVGGGIAAVQDGGQGVSSSTSTEDPVVFSRNTAVFPGNVARVGGGIYSGNVAF 300

QY 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGYDGGAI FCKNGAQ-AGSNNSGSVSPD 359  
DB 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGYDGGAI FCKNGAQ-AGSNNSGSVSPD 360

QY 360 GEGVVPFSSNVAAGKGAIIYAKKLSVANGCPVQFLNIAANDGGAIIYLGSGBELSLADYG 419  
DB 361 GEGVVPFSSNVAAGKGAIIYAKKLSVANGCPVQFLNIAANDGGAIIYLGSGBELSLADYG 420

QY 420 DIIPDGNLKRXTAKENAADVNGVTVSSQAI SMGSGGKIITTLRAKAGHOIILFNDPIEMANGN 479  
DB 421 DIIPDGNLKRXTAKENAADVNGVTVSSQAI SMGSGGKIITTLRAKAGHOIILFNDPIEMANGN 480

QY 480 NQPAQSSKLLKINDGEGYTDIVFANGSSTLVQNTIEQIRIVLREKAKLSVNSLSQTCG 539  
DB 481 NQPAQSSKLLKINDGEGYTDIVFANGSSTLVQNTIEQIRIVLREKAKLSVNSLSQTCG 540

QY 540 SLYMEAGSTWDFVTPOPPQPPAANQLITLSNLHLSLSILLANNAVTPNTPPAQDSHP 599  
DB 541 SLYMEAGSTWDFVTPOPPQPPAANQLITLSNLHLSLSILLANNAVTPNTPPAQDSHP 600

QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSGNOKINVLKQLGTQPPANAPSDLT 659  
DB 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSGNOKINVLKQLGTQPPANAPSDLT 660

QY 660 GNEMPKYQGSWKLAWDPNTANGPYTLKATWTKTYGYNPGERVASLVPNSLWGSILDI 719  
DB 661 GNEMPKYQGSWKLAWDPNTANGPYTLKATWTKTYGYNPGERVASLVPNSLWGSILDI 720

QY 720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFVYHDRDALGQGVRYISGGYSLGANSYFGSSMF 779

```
Db 721 RSAHSAIQASVDGRSYCRGLWVGSNFFYHDDRDLAQGYRIISGYSLGANSYFGSSMF 780
Qy 780 GLAFTEVFGRSKDYVVCRRNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 839
Db 781 GLAFTEVFGRSKDYVVCRRNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 840
Qy 840 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQARFSAADHESPTREGD 899
Db 841 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQARFSAADHESPTREGD 900
Qy 900 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTETLLSHQETW 959
Db 901 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTETLLSHQETW 960
Qy 960 TTDAPFLARHGVIVRGSMYASLTSNIEVYHGRYEYRDSRGYGLSAGSKVRP 1012
Db 961 TTDAPFLARHGVIVRGSMYASLTSNIEVYHGRYEYRDSRGYGLSAGSKVRP 1013

RESULT 11
US-10-498-327-1
; Sequence 1, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-1
```

```
Query Match 97.2%; Score 5121.5; DB 5; Length 1013;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 978; Conservative 23; Mismatches 11; Indels 1; Gaps 1;

Qy 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP 60
Db 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP 60

Qy 61 SAGBLTLKNLSDNSIAALPLSCFNLGSPFTVLGRGHSITPENIRTSNNGAALSADGL 120
Db 61 SAGBLTLKNLSDNSIAALPLSCFNLGSPFTVLGRGHSITPENIRTSNNGAALSADGL 120

Qy 121 FTISGFKELSPNCNLSLAVLPAATNNGSGTPTTSPSNGTIYSKTDILLANKEPSF 180
Db 121 FTISGFKELSPNCNLSLAVLPAATNNGSGTPTTSPSNGTIYSKTDILLANKEPSF 180

Qy 181 YSNLVSVDGGDAIDAKSLTVQGISKLCVQFQNTAQDGGACQVTSFSAWANEAPAFVAN 240
Db 181 YSNLVSVDGGDAIDAKSLTVQGISKLCVQFQNTAQDGGACQVTSFSAWANEAPAFVAN 240

Qy 241 VAGVGGGIAAVQDQGGQVSSSTSTEDPVVFSRNTAVEFGNVARVGGGIIYSVGNVAF 300
Db 241 VAGVGGGIAAVQDQGGQVSSSTSTEDPVVFSRNTAVEFGNVARVGGGIIYSVGNVAF 300

Qy 301 NNGKTLFLNNVASPVYLAKEQPTNGQASNTSDNTGDDGAIKCKNGAQ-AGSNNGSVSFD 359
Db 301 NNGKTLFLNNVASPVYLAKEQPTNGQASNTSDNTGDDGAIKCKNGAQ-AGSNNGSVSFD 360

Qy 360 GEGVVFSSNVAAGKGGAIYAKLISVANCQPVQFLRNIAINDGGAIIYLGESGELSADYG 419
Db 361 GEGVVFSSNVAAGKGGAIYAKLISVANCQPVQFLRNIAINDGGAIIYLGESGELSADYG 420
```

```
Qy 420 DIIIFDGNLKRRTAKENAADVNGVTVSSQAIISGSGGKITTLRAKAGHQILLFNDPIEMANGN 479
Db 421 DIIIFDGNLKRRTAKENAADVNGVTVSSQAIISGSGGKITTLRAKAGHQILLFNDPIEMANGN 480
Qy 480 NOPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLRBKAKLSVNSISQTTGG 539
Db 481 NOPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLRBKAKLSVNSISQTTGG 540
Qy 540 SLYNEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLIANNAVTPPTNPPAQDSHP 599
Db 541 SLYNEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLIANNAVTPPTNPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIFFEDLDLDTAYDRYDNLGSKNOKINVLKQLGTYKPPANAPSIDL 659
Db 601 AVIGSTTAGSVTISGPIFFEDLDLDTAYDRYDNLGSKNOKIDVLKQLGTYKPPANAPSIDL 660
Qy 660 GNEPMKYGYGSGWKLAWDPNTANNPPTLKATWTKTCYVPCGPVRSIVPNSLWGSILDI 719
Db 661 GNEPMKYGYGSGWKLAWDPNTANNPPTLKATWTKTCYVPCGPVRSIVPNSLWGSILDI 720
Qy 720 RSAHSAIQASVDGRSYCRGLWVGSNFFYHDDRDLAQGYRIISGYSLGANSYFGSSMF 779
Db 721 RSAHSAIQASVDGRSYCRGLWVGSNFFYHDDRDLAQGYRIISGYSLGANSYFGSSMF 780
Qy 780 GLAFTEVFGRSKDYVVCRRNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 839
Db 781 GLAFTEVFGRSKDYVVCRRNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 840
Qy 840 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQARFSAADHESPTREGD 899
Db 841 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQARFSAADHESPTREGD 900
Qy 900 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTETLLSHQETW 959
Db 901 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTETLLSHQETW 960
Qy 960 TTDAPFLARHGVIVRGSMYASLTSNIEVYHGRYEYRDSRGYGLSAGSKVRP 1012
Db 961 TTDAPFLARHGVIVRGSMYASLTSNIEVYHGRYEYRDSRGYGLSAGSKVRP 1013

RESULT 12
US-10-498-327-3
; Sequence 3, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-3
```

```
Query Match 97.2%; Score 5121.5; DB 5; Length 1013;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 978; Conservative 23; Mismatches 11; Indels 1; Gaps 1;

Qy 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP 60
Db 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP 60

Qy 61 SAGBLTLKNLSDNSIAALPLSCFNLGSPFTVLGRGHSITPENIRTSNNGAALSADGL 120
Db 61 SAGBLTLKNLSDNSIAALPLSCFNLGSPFTVLGRGHSITPENIRTSNNGAALSADGL 120
```

121 FTIEGKELSPNCSNLLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180  
121 FTIEGKELSPNCSNLLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180  
181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTVTSFSAANEAPAFVAN 240  
181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTVTSFSAANEAPAFVAN 240  
241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVPSRNTAVFEDGNVARVGGGIYSYGNVAPL 300  
241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVPSRNTAVFEDGNVARVGGGIYSYGNVAPL 300  
301 NNGKTLFLNNVSPVIAAKQTSQASNTSNNGYDGGAIKCKNGAQ-AGSNNSGSVSD 359  
301 NNGKTLFLNNVSPVIAAKQTSQASNTSNNGYDGGAIKCKNGAQ-AGSNNSGSVSD 360  
360 GEGVVPFSSNVAAGKGAIYAKKLSVANGCPVQFLRNANDGAIYLGESGELSADY 419  
361 GEGVVPFSSNVAAGKGAIYAKKLSVANGCPVQFLRNANDGAIYLGESGELSADY 420  
420 DIIFDGNLARTAKENAADVNGTVSSQAI SMGSGGKITTLLRAKAGHQILFNDPIEMANG 479  
421 DIIFDGNLARTAKENAADVNGTVSSQAI SMGSGGKITTLLRAKAGHQILFNDPIEMANG 480  
480 NPAQSGKLLKINDGEGYTGDIVPANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSOTGG 539  
481 NPAQSGSBLKINDGEGYTGDIVPANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSOTGG 540  
540 SLYMEAGSTWDTVPPOPPQAPNOLITLNLHLSLLANNATVNPPTPPAQDSHP 599  
541 SLYMEAGSTWDTVPPOPPQAPNOLITLNLHLSLLANNATVNPPTPPAQDSHP 600  
600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKIDVLKQLGTQPSANAPSDTL 659  
601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKIDVLKQLGTQPSANAPSDTL 660  
660 GNEMPKYQGSWKLAWDNTANGPYTLKATWTKTGNPGRVPSVLSVNSLWGSILDI 719  
661 GNEMPKYQGSWKLAWDNTANGPYTLKATWTKTGNPGRVPSVLSVNSLWGSILDI 720  
720 RSAHSAIQASVGRSVCRLWVSGVSNFFVHDDALGQGVYISGGVSLGANSYFGSSMF 779  
721 RSAHSAIQASVGRSVCRLWVSGVSNFFVHDDALGQGVYISGGVSLGANSYFGSSMF 780  
780 GLAFTVFGRSKDYVCRSNHACIGSVYLSSTQALCGSYLFGDAFLRASVYFGNQHMKT 839  
781 GLAFTVFGRSKDYVCRSNHACIGSVYLSSTQALCGSYLFGDAFLRASVYFGNQHMKT 840  
840 SYTFABESDVRDNNCLAGBIGAGLPVITPSPKLYLNELRPFVQABFSYADHESFTEGD 899  
841 SYTFABESDVRDNNCLAGBIGAGLPVITPSPKLYLNELRPFVQABFSYADHESFTEGD 900  
900 QABAPKSGHLLNLVSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTTSGTETLLSHQETW 959  
901 QABAPKSGHLLNLVSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTTSGTETLLSHQETW 960  
960 TTDAPFLARHGVIIVRGSMYASLTSNIEVYGHGRYEDRSDAGSGLSAGSVRF 1012  
961 TTDAPFLARHGVIIVRGSMYASLTSNIEVYGHGRYEDRSDAGSGLSAGSVRF 1013

RESULT 13  
US-10-701-844-43  
; Sequence 43, Application US/10701844  
; Publication No. US20040067524A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/10701,844  
; CURRENT FILING DATE: 2003-11-04

; PRIOR APPLICATION NUMBER: US/09/612,402B  
; PRIOR FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 43  
; LENGTH: 984  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-701-844-43

Query Match 97.1%; Score 5116; DB 4; Length 984;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 EIMVPGIYDGETLTVSPFYTVIGDPSTTVPSAGELTLKLNLDNSIALPLSCFNGLLGS 88  
DB 1 EIMVPGIYDGETLTVSPFYTVIGDPSTTVPSAGELTLKLNLDNSIALPLSCFNGLLGS 60  
QY 89 FTVLGRGHSILPENIRTSNGAALSADGLFTTIEGFKELSFNSCNSLLAVLPAATTNK 148  
DB 61 FTVLGRGHSILPENIRTSNGAALSADGLFTTIEGFKELSFNSCNSLLAVLPAATTNK 120  
QY 149 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCV 208  
DB 121 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCV 180  
QY 209 QNTAODGACQVTVTSFSAANEAPAFVANVAGVGGIAAQQDQGGVSSSTSTEDP 268  
DB 181 QNTAODGACQVTVTSFSAANEAPAFVANVAGVGGIAAQQDQGGVSSSTSTEDP 240  
QY 269 VVPSRNTAVFEDGNVARVGGGIYSYGNVAFNLNGKTLFLNNVSPVYIAAKQTSQAS 328  
DB 241 VVPSRNTAVFEDGNVARVGGGIYSYGNVAFNLNGKTLFLNNVSPVYIAAKQTSQAS 300  
QY 329 NTSNNYDGGAIKCKNGAQAGSNNSGSVFDGEGVVFSSNVAAGKGAIYAKKLSVANC 388  
DB 301 NTSNNYDGGAIKCKNGAQAGSNNSGSVFDGEGVVFSSNVAAGKGAIYAKKLSVANC 360  
QY 389 GPVQFLRNANDGAIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGTVSSQAI 448  
DB 361 GPVQFLRNANDGAIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGTVSSQAI 420  
QY 449 SMGSGGKITTLLRAKAGHQILFNDPIEMANGNQPAOSSKLLKINDGEGYTGDIVPANGSS 508  
DB 421 SMGSGGKITTLLRAKAGHQILFNDPIEMANGNQPAOSSKLLKINDGEGYTGDIVPANGSS 480  
QY 509 TLYQNTVIEQGRIVLREKAKLSVNSLSQTSGLYMEAGSTWDFVTPPQPPAANQLIT 568  
DB 481 TLYQNTVIEQGRIVLREKAKLSVNSLSQTSGLYMEAGSTWDFVTPPQPPAANQLIT 540  
QY 569 LSNLHLSLSLLANNATVNPPTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDR 628  
DB 541 LSNLHLSLSLLANNATVNPPTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDR 600  
QY 629 YDWLGSNQKINVLKQLGTQPPANAPSDTLTGNEPKYQGSWKLAWDNTANGPYTL 688  
DB 601 YDWLGSNQKINVLKQLGTQPPANAPSDTLTGNEPKYQGSWKLAWDNTANGPYTL 660  
QY 689 KATWTKTGNPGRVPSVLSVNSLWGSILDIRSAHSAIQASVGRSVCRLWVSGVSNFF 748  
DB 661 KATWTKTGNPGRVPSVLSVNSLWGSILDIRSAHSAIQASVGRSVCRLWVSGVSNFF 720  
QY 749 YHDDRDLGQGVYISGGVSLGANSYFGSSMFLAFTVFGRSKDYVCRSNHACIGSVY 808  
DB 721 YHDDRDLGQGVYISGGVSLGANSYFGSSMFLAFTVFGRSKDYVCRSNHACIGSVY 780  
QY 809 LSTQALCGSYLFGDAFLRASVYFGNQHMKTSTFAEESDVRDNNCLAGBIGAGLPV 868  
DB 781 LSTQALCGSYLFGDAFLRASVYFGNQHMKTSTFAEESDVRDNNCLAGBIGAGLPV 840  
QY 869 TPSKLYLNELRPFVQABFSYADHESFTEGDQAPAPKSGHLLNLVSPVGVKFCRCSSTHP 928

Db 841 TPSKLYLNELRPVQAEFSYADHESFTTEGQARAFKSGHLNLNLSVPVGVKFDRCSSTHP 900  
Qy 929 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVY 988  
Db 901 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVY 960  
Qy 989 GHGREYRDASRGYGLSAGSRVP 1012  
Db 961 GHGREYRDASRGYGLSAGSRVP 984

RESULT 14  
US-10-766-711-43  
; Sequence 43, Application US/10766711  
; Publication No. US20040137005A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-062-999  
; CURRENT APPLICATION NUMBER: US/10/766.711  
; CURRENT FILING DATE: 2004-01-27  
; PRIOR APPLICATION NUMBER: US/08/942.596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 43  
; LENGTH: 984  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-766-711-43

Query Match 97.1%; Score 5116; DB 4; Length 984;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFVSAGELTLKLNLSIAALPLSCFNLGGS 88  
Db 1 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFVSAGELTLKLNLSIAALPLSCFNLGGS 60

Qy 89 FTVLGRGHSFTFNIRSTNGAALSNAADGLFTIEGFKELSPNCNLSLAVLPAATNK 148  
Db 61 FTVLGRGHSFTFNIRSTNGAALSNAADGLFTIEGFKELSPNCNLSLAVLPAATNK 120

Qy 149 GSQPTTTTSPNGTIYSKTDLLLNNEKFPBFYGNLVSGDGAIDAKSLTVQGISKLCVP 208  
Db 121 GSQPTTTTSPNGTIYSKTDLLLNNEKFPBFYGNLVSGDGAIDAKSLTVQGISKLCVP 180

Qy 209 QENTAQADGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVDQGGQVSSSTSTEDP 268  
Db 181 QENTAQADGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVDQGGQVSSSTSTEDP 240

Qy 269 VVSFRTAVSFEDGNVAVRGGIYSYGNVAFNLNKGTLFLNNVASPVVIAAKOPTSGOAS 328  
Db 241 VVSFRTAVSFEDGNVAVRGGIYSYGNVAFNLNKGTLFLNNVASPVVIAAKOPTSGOAS 300

Qy 329 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDPGEGVFPFSSNVAAGKGAIIYAKKLSVANC 388  
Db 301 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDPGEGVFPFSSNVAAGKGAIIYAKKLSVANC 360

Qy 389 GPVQFLRNINDGAIYLGESGELSLSDYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 448  
Db 361 GPVQFLRNINDGAIYLGESGELSLSDYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 420

Qy 449 SMGSGGKITTLURAKAGHQILFNDPIEMANGNQPAQSKLLKINDGGYTGDIIVFANGSS 508  
Db 421 SMGSGGKITTLURAKAGHQILFNDPIEMANGNQPAQSKLLKINDGGYTGDIIVFANGSS 480

Qy 509 TLYQNVITBQRIIVLREKAKLSVNSLSQTGGSLSYMEAGSTWDFVTPPQPPQPPAANQLIT 568  
Db 481 TLYQNVITBQRIIVLREKAKLSVNSLSQTGGSLSYMEAGSTWDFVTPPQPPQPPAANQLIT 540

Qy 569 LSNLHLSLSLLANNVNTNPTTPAODSHPAVIGSTTAGSVTISGPIFPEDLDDTAYDR 628  
Db 541 LSNLHLSLSLLANNVNTNPTTPAODSHPAVIGSTTAGSVTISGPIFPEDLDDTAYDR 600

Qy 629 YDMLGSNOKNLVNLKQLGTGTPPANAPSDLTGLNEMPKYGYGSKWLAWDPNTANNPPTL 688  
Db 601 YDMLGSNOKNLVNLKQLGTGTPPANAPSDLTGLNEMPKYGYGSKWLAWDPNTANNPPTL 660

Qy 689 KATWTKGTYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNFF 748  
Db 661 KATWTKGTYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNFF 720

Qy 749 YHDERDALGQGYRYTISGGYSLGANSYFGSSNPFGLAFTVFRGSKDYVYVCRSNHACISGVY 808  
Db 721 YHDERDALGQGYRYTISGGYSLGANSYFGSSNPFGLAFTVFRGSKDYVYVCRSNHACISGVY 780

Qy 809 LSTQOALCGSYLFGDAFIRASYGFGNOMKTSYTFPAGESDVRWNNCLAGEIGAGLPIVI 868  
Db 781 LSTQOALCGSYLFGDAFIRASYGFGNOMKTSYTFPAGESDVRWNNCLAGEIGAGLPIVI 840

Qy 869 TPSKLYLNELRPVQAEFSYADHESFTTEGQARAFKSGHLNLNLSVPVGVKFDRCSSTHP 928  
Db 841 TPSKLYLNELRPVQAEFSYADHESFTTEGQARAFKSGHLNLNLSVPVGVKFDRCSSTHP 900

Qy 929 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVY 988  
Db 901 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVY 960

Qy 989 GHGREYRDASRGYGLSAGSRVP 1012  
Db 961 GHGREYRDASRGYGLSAGSRVP 984

RESULT 15  
US-10-931-779-43  
; Sequence 43, Application US/10931779  
; Publication No. US20050048557A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: BP104  
; CURRENT APPLICATION NUMBER: US/10/931.779  
; CURRENT FILING DATE: 2004-09-01  
; PRIOR APPLICATION NUMBER: 09/542,520  
; PRIOR FILING DATE: 200-04-03  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 43  
; LENGTH: 984  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-931-779-43

Query Match 97.1%; Score 5116; DB 5; Length 984;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFVSAGELTLKLNLSIAALPLSCFNLGGS 88  
Db 1 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFVSAGELTLKLNLSIAALPLSCFNLGGS 60

Qy 89 FTVLGRGHSFTFNIRSTNGAALSNAADGLFTIEGFKELSPNCNLSLAVLPAATNK 148  
Db 61 FTVLGRGHSFTFNIRSTNGAALSNAADGLFTIEGFKELSPNCNLSLAVLPAATNK 120

Qy 149 GSQPTTTTSPNGTIYSKTDLLLNNEKFPBFYGNLVSGDGAIDAKSLTVQGISKLCVP 208  
Db 121 GSQPTTTTSPNGTIYSKTDLLLNNEKFPBFYGNLVSGDGAIDAKSLTVQGISKLCVP 180

Qy 209 QENTAQADGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVDQGGQVSSSTSTEDP 268  
Db 181 QENTAQADGGACQVTVSFSAMANEAPAFVANVAGVRGGIAAVDQGGQVSSSTSTEDP 240

269 VVSFRTAVRFDGNAVUGGIIYSGNVAFNNKTLFLANNVSPVIAAKOPTSGOAS 328  
Db VVSFRTAVRFDGNAVUGGIIYSGNVAFNNKTLFLANNVSPVIAAKOPTSGOAS 300  
329 NTSNNYDGGGAIKCKNGAQAAGNNSSGVSFDGEGVVFSSNVAAGKGAIAKAKLSVANC 388  
Db NTSNNYDGGGAIKCKNGAQAAGNNSSGVSFDGEGVVFSSNVAAGKGAIAKAKLSVANC 360  
389 GPVQFLRNANDGGAIIYSGSGLSADYGDIIIFDGNLKRTAKENAADVNGVTVSSQAI 448  
Db GPVQFLRNANDGGAIIYSGSGLSADYGDIIIFDGNLKRTAKENAADVNGVTVSSQAI 420  
449 SMGSGGKIITIRAKAGHOITLNDPILFEMANGNNOQAQSSKLLKINDGEGVTGDIVANGSS 508  
Db SMGSGGKIITIRAKAGHOITLNDPILFEMANGNNOQAQSSKLLKINDGEGVTGDIVANGSS 480  
509 TLYQNVTTIEQRIVLREKAKLSVNSLSQGGSLYMEAGTMDFTVPPQPPAANQLIT 568  
Db TLYQNVTTIEQRIVLREKAKLSVNSLSQGGSLYMEAGTMDFTVPPQPPAANQLIT 540  
569 LSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIPFEDLDDTAYDR 628  
Db LSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIPFEDLDDTAYDR 600  
629 YDWLGSNQKINVLKQLGTHKPPANAPSDTLTGNEMPKYQYQGSWKLAWDPNTANNPYTL 688  
Db YDWLGSNQKINVLKQLGTHKPPANAPSDTLTGNEMPKYQYQGSWKLAWDPNTANNPYTL 660  
689 KATWTKTGNPGERVASLVPNSLWMSIILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFF 748  
Db KATWTKTGNPGERVASLVPNSLWMSIILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFF 720  
749 YHDRDALGQYRIYSGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRSNNHACIGSVY 808  
Db YHDRDALGQYRIYSGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRSNNHACIGSVY 780  
809 LSTQQAALCSYLFQDAFIASVFGNQHMKTSTTPAESDVRDNNCLAGEIGAGLPIVI 868  
Db LSTQQAALCSYLFQDAFIASVFGNQHMKTSTTPAESDVRDNNCLAGEIGAGLPIVI 840  
869 TPKSKYLNELRPPVQAEFSVADHESFTEEGDQARAFKSHLLNLSVPVGVKFCRCSSTHP 928  
Db TPKSKYLNELRPPVQAEFSVADHESFTEEGDQARAFKSHLLNLSVPVGVKFCRCSSTHP 900  
929 NKYSFMAAICDAYRTISGTETLLSHQETWTTTDAFLARHGYYVVRGSMYASLTSNIEVY 988  
Db NKYSFMAAICDAYRTISGTETLLSHQETWTTTDAFLARHGYYVVRGSMYASLTSNIEVY 960  
989 GHGRYERDASRGYGLSAGSRVRF 1012  
Db GHGRYERDASRGYGLSAGSRVRF 984

RESULT 16  
US-10-498-327-17  
; Sequence 17, Application US/10498327  
; Publication No. US20050106162A1  
; GENERAL INFORMATION:  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis  
; FILE REFERENCES: 002441.00085  
; CURRENT APPLICATION NUMBER: US/10/498,327  
; CURRENT FILING DATE: 2004-06-10  
; PRIOR APPLICATION NUMBER: PCT/IB02/05761  
; PRIOR FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 262  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis

US-10-498-327-17

Query Match 97.1%; Score 5112.5; DB 5; Length 1013;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 977; Conservative 23; Mismatches 12; Indels 1; Gaps 1;  
QY 1 MQTSFHKFPFLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFVTVIGDPGGTTVP 60  
Db 1 MQTSFHKFPFLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFVTVIGDPGGTTVP 60  
61 SAGELTLKVLNDSNLAALPLSCFGLNLSFTVLGRHSITFEENIRTSNMGALSNSAADGL 120  
QY 61 SAGELTLKVLNDSNLAALPLSCFGLNLSFTVLGRHSITFEENIRTSNMGALSNSAADGL 120  
Db 61 SAGELTLKVLNDSNLAALPLSCFGLNLSFTVLGRHSITFEENIRTSNMGALSNSAADGL 120  
121 FTIEGFKELSFNCNLSLAVLPAATNKGSTPTPTTSTPSNGTIYKTDLLLNNEKPSF 180  
QY 121 FTIEGFKELSFNCNLSLAVLPAATNKGSTPTPTTSTPSNGTIYKTDLLLNNEKPSF 180  
Db 121 FTIEGFKELSFNCNLSLAVLPAATNKGSTPTPTTSTPSNGTIYKTDLLLNNEKPSF 180  
181 YSNLVSGDGGAIKAKSLTVQGISKLCPQENTAQADGGACQVVTFSFMANEAPAFVAN 240  
QY 181 YSNLVSGDGGAIKAKSLTVQGISKLCPQENTAQADGGACQVVTFSFMANEAPAFVAN 240  
Db 181 YSNLVSGDGGAIKAKSLTVQGISKLCPQENTAQADGGACQVVTFSFMANEAPAFVAN 240  
241 VAGVRGGGIAAVQDQOQGVSSSTSTEDPVVSFRNTAVEFDGNVARVGGIYSGNVAF 300  
QY 241 VAGVRGGGIAAVQDQOQGVSSSTSTEDPVVSFRNTAVEFDGNVARVGGIYSGNVAF 300  
Db 241 VAGVRGGGIAAVQDQOQGVSSSTSTEDPVVSFRNTAVEFDGNVARVGGIYSGNVAF 300  
301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNYDGGAIKCKNGAQ-AGSNNSVSFSD 359  
QY 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNYDGGAIKCKNGAQ-AGSNNSVSFSD 359  
Db 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNYDGGAIKCKNGAQ-AGSNNSVSFSD 359  
360 GEGVVFSSNVAAGKGAIAKLSVANCQVQFLRNANDGGAIIYLGESGSLSADY 419  
QY 360 GEGVVFSSNVAAGKGAIAKLSVANCQVQFLRNANDGGAIIYLGESGSLSADY 419  
Db 360 GEGVVFSSNVAAGKGAIAKLSVANCQVQFLRNANDGGAIIYLGESGSLSADY 419  
420 DIIFDGNLKRTAKENAADVNGVTVSSQAIKCKNGAQAAGNNSSGVSFD 360  
QY 420 DIIFDGNLKRTAKENAADVNGVTVSSQAIKCKNGAQAAGNNSSGVSFD 360  
Db 420 DIIFDGNLKRTAKENAADVNGVTVSSQAIKCKNGAQAAGNNSSGVSFD 360  
480 NQPAOSSKLLKINDGEGTGDIVFANGSTLQNTVIEQRIVLREKAKLSVNSLSQGG 539  
QY 480 NQPAOSSKLLKINDGEGTGDIVFANGSTLQNTVIEQRIVLREKAKLSVNSLSQGG 539  
Db 480 NQPAOSSKLLKINDGEGTGDIVFANGSTLQNTVIEQRIVLREKAKLSVNSLSQGG 539  
540 SLYMEAGSTFDPVTPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599  
QY 540 SLYMEAGSTFDPVTPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599  
Db 540 SLYMEAGSTFDPVTPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599  
600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDWLGSNQKINVLKQLGTHKPPANAPSDTL 659  
QY 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDWLGSNQKINVLKQLGTHKPPANAPSDTL 659  
Db 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDWLGSNQKINVLKQLGTHKPPANAPSDTL 659  
660 GNEMPKYQYQGSWKLAWDPNTANNPYTLKATWTKTGNPGERVASLVPNSLWMSIILDI 719  
QY 660 GNEMPKYQYQGSWKLAWDPNTANNPYTLKATWTKTGNPGERVASLVPNSLWMSIILDI 719  
Db 660 GNEMPKYQYQGSWKLAWDPNTANNPYTLKATWTKTGNPGERVASLVPNSLWMSIILDI 719  
720 RSHASAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRIYSGYSLGANSYFGSSMF 779  
QY 720 RSHASAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRIYSGYSLGANSYFGSSMF 779  
Db 720 RSHASAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRIYSGYSLGANSYFGSSMF 779  
780 GLAFTEVFGRSKDYVVCRSNNHACIGSVYLSQQAALCSYLFQDAFIASVFGNQHMKT 839  
QY 780 GLAFTEVFGRSKDYVVCRSNNHACIGSVYLSQQAALCSYLFQDAFIASVFGNQHMKT 839  
Db 780 GLAFTEVFGRSKDYVVCRSNNHACIGSVYLSQQAALCSYLFQDAFIASVFGNQHMKT 839  
840 SYTFAESDVRDNNCLAGEIGAGLPIVITPSKLYLNELRPPVQAEFSVADHESFTEEGD 899  
QY 840 SYTFAESDVRDNNCLAGEIGAGLPIVITPSKLYLNELRPPVQAEFSVADHESFTEEGD 899  
Db 840 SYTFAESDVRDNNCLAGEIGAGLPIVITPSKLYLNELRPPVQAEFSVADHESFTEEGD 899  
900 QARAFKSHLLNLSVPVGVKFCRCSSTHPNKSFMAAICDAYRTISGTETLLSHQETW 959  
QY 900 QARAFKSHLLNLSVPVGVKFCRCSSTHPNKSFMAAICDAYRTISGTETLLSHQETW 959  
Db 900 QARAFKSHLLNLSVPVGVKFCRCSSTHPNKSFMAAICDAYRTISGTETLLSHQETW 959  
960 TTDAFLARHGYYVVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012  
QY 960 TTDAFLARHGYYVVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012  
Db 960 TTDAFLARHGYYVVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012



```

RESULT 17
US-10-701-844-15
; Sequence 15, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-15

Query Match      97.0%; Score 5106.5; DB 4; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;

Qy      1  MQTSPHFKPFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFVTYIGDPSGTTVP 60
Db      1  MQTSPHFKPFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFVTYIGDPSGTTVP 60

Qy      61  SAGELTLKNDLSIAALPLSCFNLGGSFTVLGRGHSITFTENIRTSNGAALSADGL 120
Db      61  SAGELTLKNDLSIAALPLSCFNLGGSFTVLGRGHSITFTENIRTSNGAALSADGL 120

Qy      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180
Db      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180

Qy      181  YSNLVSDGGGAIKAKLSVANGCPVQFLRNTANDGGAIYLGESGELSADYG 419
Db      181  YSNLVSDGGGAIKAKLSVANGCPVQFLRNTANDGGAIYLGESGELSADYG 419

Qy      420  DIIIDGNLKRITAKENADVNGTVSSQAIISMGSGKITTLLAKAGHQTLLFNDPIEMANGN 479
Db      420  DIIIDGNLKRITAKENADVNGTVSSQAIISMGSGKITTLLAKAGHQTLLFNDPIEMANGN 479

Qy      480  NPAQSSKLLKINDGEGTVGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGG 539
Db      480  NPAQSSKLLKINDGEGTVGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGG 539

Qy      540  SLYNEAGSTWDFVTPQPPQAPPAANOLITLNLHLISLSSLLANNAVTPNPPPAQDSHP 599
Db      540  SLYNEAGSTWDFVTPQPPQAPPAANOLITLNLHLISLSSLLANNAVTPNPPPAQDSHP 599

Qy      600  AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGNSQKINYLKLQGLGTPPANAPSIDLTL 659
Db      600  AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGNSQKIDVLKLQGLGTPPANAPSIDLTL 660

```

```

Qy      660  GNEMPKYCGSKWLANDPNTANNPGYTLAKATWTKTGYNCPERVASLVPNSLWGSILDI 719
Db      660  GNEMPKYCGSKWLANDPNTANNPGYTLAKATWTKTGYNCPERVASLVPNSLWGSILDI 720

Qy      720  RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHRRDALQGGYRIISGGYSLGANSYFGSSMF 779
Db      720  RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHRRDALQGGYRIISGGYSLGANSYFGSSMF 780

Qy      780  GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTOQALCGSYLFGDAFIRASYGFGNQHMKT 839
Db      780  GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTOQALCGSYLFGDAFIRASYGFGNQHMKT 840

Qy      840  SYTFAEESDVWNNCLAGHIGAGLPVITPSPKLYINELRPPFVQAESYADHESFTBECD 899
Db      840  SYTFAEESDVWNNCLAGHIGAGLPVITPSPKLYINELRPPFVQAESYADHESFTBECD 900

Qy      900  QARAFKSHLNLNLSVPVGVKFDRCSSHPNKYSFMAAYICDAYRTISGTETTLSSHORTW 959
Db      900  QARAFKSHLNLNLSVPVGVKFDRCSSHPNKYSFMAAYICDAYRTISGTETTLSSHORTW 960

Qy      960  TTDAPHLARHGVVVRGSMYASLTNSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012
Db      960  TTDAPHLARHGVVVRGSMYASLTNSNIEVYGHGRYERDASRGYGLSAGSRVRF 1013

RESULT 18
US-10-701-844-16
; Sequence 16, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-16

Query Match      97.0%; Score 5106.5; DB 4; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

Qy      1  MQTSPHFKPFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFVTYIGDPSGTTVP 60
Db      1  MQTSPHFKPFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFVTYIGDPSGTTVP 60

Qy      61  SAGELTLKNDLSIAALPLSCFNLGGSFTVLGRGHSITFTENIRTSNGAALSADGL 120
Db      61  SAGELTLKNDLSIAALPLSCFNLGGSFTVLGRGHSITFTENIRTSNGAALSADGL 120

Qy      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180
Db      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180

Qy      181  YSNLVSDGGGAIKAKLSVANGCPVQFLRNTANDGGAIYLGESGELSADYG 240
Db      181  YSNLVSDGGGAIKAKLSVANGCPVQFLRNTANDGGAIYLGESGELSADYG 240

Qy      241  VAGVGGGIAAVQDQGGVSSSTSTEDPVVSFSRNTAVEFDGNVARGGGIYSYGNVAF 300
Db      241  VAGVGGGIAAVQDQGGVSSSTSTEDPVVSFSRNTAVEFDGNVARGGGIYSYGNVAF 300

Qy      301  NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNNTYDGGAIKFKNGAQ-AGSNNSGSVSFD 359
Db      301  NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNNTYDGGAIKFKNGAQ-AGSNNSGSVSFD 360

Qy      360  GEGVVFSSNVAAGGGGAIYAKLSVANGCPVQFLRNTANDGGAIYLGESGELSADYG 419
Db      360  GEGVVFSSNVAAGGGGAIYAKLSVANGCPVQFLRNTANDGGAIYLGESGELSADYG 420

Qy      420  DIIIDGNLKRITAKENADVNGTVSSQAIISMGSGKITTLLAKAGHQTLLFNDPIEMANGN 479
Db      420  DIIIDGNLKRITAKENADVNGTVSSQAIISMGSGKITTLLAKAGHQTLLFNDPIEMANGN 480

Qy      480  NPAQSSKLLKINDGEGTVGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGG 539
Db      480  NPAQSSKLLKINDGEGTVGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGG 540

Qy      540  SLYNEAGSTWDFVTPQPPQAPPAANOLITLNLHLISLSSLLANNAVTPNPPPAQDSHP 599
Db      540  SLYNEAGSTWDFVTPQPPQAPPAANOLITLNLHLISLSSLLANNAVTPNPPPAQDSHP 600

Qy      600  AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGNSQKINYLKLQGLGTPPANAPSIDLTL 659
Db      600  AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGNSQKIDVLKLQGLGTPPANAPSIDLTL 660

```

Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAAGSNNSGSVSFD 360  
Qy 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGAIIYLGESGELSADYG 419  
Db 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGAIIYLGESGELSADYG 420  
Qy 420 DIIFGNLRKTAKENADVGVTVSSQAI SMGSGGKIITTLRAKAGHOIILFNDPIEMANGN 479  
Db 421 DIIFGNLRKTAKENADVGVTVSSQAI SMGSGGKIITTLRAKAGHOIILFNDPIEMANGN 480  
Qy 480 NPOASSKLLKINDGEGYTGDI VFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 539  
Db 481 NPOASSSEPLKINDGEGYTGDI VFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 540  
Qy 540 SLYMEAGSTWDFVTPOPPQPPAANQLITLSNHLSSLANNVTPNPPPAQDSHP 599  
Db 541 SLYMEAGSTLDFVTPOPPQPPAANQLITLSNHLSSLANNVTPNPPPAQDSHP 600  
Qy 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDTL 659  
Db 601 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDTL 660  
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKTGYNPGERVASLVPNSLWGSILDI 719  
Db 661 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKTGYNPGERVASLVPNSLWGSILDI 720  
Qy 720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLGQGYRYISGGYSLGANSYFGSSMF 779  
Db 721 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLGQGYRYISGGYSLGANSYFGSSMF 780  
Qy 780 GLAFTVFGRSKDYVYVCRSNHHCIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 839  
Db 781 GLAFTVFGRSKDYVYVCRSNHHCIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 840  
Qy 840 SYTFABESDVRWNNCLVGEIGVGLPIVITPSPKLYLNELRPPFVQABFSYADHESFTEEG 899  
Db 841 SYTFABESDVRWNNCLVGEIGVGLPIVITPSPKLYLNELRPPFVQABFSYADHESFTEEG 900  
Qy 900 QARAFKSGHLLNLISVPVGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 959  
Db 901 QARAFKSGHLLNLISVPVGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 960  
Qy 960 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGREYEDRSRGYGLSAGSVRFP 1012  
Db 961 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGREYEDRSRGYGLSAGSVRFP 1013

RESULT 19  
US-10-766-711-15  
; Sequence 15, Application US/10766711  
; Publication No. US20040137005A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-062-999  
; CURRENT APPLICATION NUMBER: US/10766,711  
; CURRENT FILING DATE: 2004-01-27  
; PRIOR APPLICATION NUMBER: US/08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 15  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-766-711-15

Query Match 97.0%; Score 5106.5; DB 4; Length 1013;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;

Qy 1 MOTSFHKFFLNLAYSCCSLNGGGYAAABIMVPOGIYDGETILTVSPFYTVIGDPSGTTFV 60  
Db 1 MOTSFHKFFLNLAYSCCSLNGGGYAAABIMVPOGIYDGETILTVSPFYTVIGDPSGTTFV 60  
Qy 61 SAGELTTLKNLONSIAALPLSCFNLGSLTFVLGRGHSITTFENIRISTNGAALSNAADGL 120  
Db 61 SAGELTTLKNLONSIAALPLSCFNLGSLTFVLGRGHSITTFENIRISTNGAALSNAADGL 120  
Qy 121 FTIEGFKELSFNSCNLSLAVLPAATTNKSGSQPTPTTSTPSNGTIYKSTDLTLLLNNEKFSF 180  
Db 121 FTIEGFKELSFNSCNLSLAVLPAATTNKSGSQPTPTTSTPSNGTIYKSTDLTLLLNNEKFSF 180  
Qy 181 YSNLVSGDGAIDAKSLTVQGISLKVQENTQAADGACQVVTSPFSAMANEAPAFVAN 240  
Db 181 YSNLVSGDGAIDAKSLTVQGISLKVQENTQAADGACQVVTSPFSAMANEAPAFVAN 240  
Qy 241 VAGVGGGIIAAVODGQGGVSSSTSTEDPVVPSRNTAVFEFGNVARVGGGIYSYGNVAF 300  
Db 241 VAGVGGGIIAAVODGQGGVSSSTSTEDPVVPSRNTAVFEFGNVARVGGGIYSYGNVAF 300  
Qy 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAAGSNNSGSVSFD 359  
Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAAGSNNSGSVSFD 360  
Qy 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGAIIYLGESGELSADYG 419  
Db 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGAIIYLGESGELSADYG 420  
Qy 420 DIIFGNLRKTAKENADVGVTVSSQAI SMGSGGKIITTLRAKAGHOIILFNDPIEMANGN 479  
Db 421 DIIFGNLRKTAKENADVGVTVSSQAI SMGSGGKIITTLRAKAGHOIILFNDPIEMANGN 480  
Qy 480 NPOASSKLLKINDGEGYTGDI VFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 539  
Db 481 NPOASSSEPLKINDGEGYTGDI VFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 540  
Qy 540 SLYMEAGSTWDFVTPOPPQPPAANQLITLSNHLSSLANNVTPNPPPAQDSHP 599  
Db 541 SLYMEAGSTLDFVTPOPPQPPAANQLITLSNHLSSLANNVTPNPPPAQDSHP 600  
Qy 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDTL 659  
Db 601 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDTL 660  
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKTGYNPGERVASLVPNSLWGSILDI 719  
Db 661 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKTGYNPGERVASLVPNSLWGSILDI 720  
Qy 720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLGQGYRYISGGYSLGANSYFGSSMF 779  
Db 721 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLGQGYRYISGGYSLGANSYFGSSMF 780  
Qy 780 GLAFTVFGRSKDYVYVCRSNHHCIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 839  
Db 781 GLAFTVFGRSKDYVYVCRSNHHCIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 840  
Qy 840 SYTFABESDVRWNNCLVGEIGVGLPIVITPSPKLYLNELRPPFVQABFSYADHESFTEEG 899  
Db 841 SYTFABESDVRWNNCLVGEIGVGLPIVITPSPKLYLNELRPPFVQABFSYADHESFTEEG 900  
Qy 900 QARAFKSGHLLNLISVPVGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 959  
Db 901 QARAFKSGHLLNLISVPVGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 960  
Qy 960 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGREYEDRSRGYGLSAGSVRFP 1012  
Db 961 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGREYEDRSRGYGLSAGSVRFP 1013

RESULT 20  
US-10-766-711-16  
; Sequence 16, Application US/10766711  
; Publication No. US20040137005A1

; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-062-999  
; CURRENT APPLICATION NUMBER: US/10/766,711  
; CURRENT FILING DATE: 2004-01-27  
; PRIOR APPLICATION NUMBER: US/08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-766-711-16

Query Match 97.0%; Score 5106.5; DB 4; Length 1013;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

Qy 1 MQTSPHKFPLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFTYVIGDPSGTTVF 60  
Db 1 MQTSPHKFPLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFTYVIGDPSGTTVF 60

Qy 61 SAGELTLKNDLSIAALPLSCFNLGSLFTVLGRGHSITPENIRTSNGAALSADGL 120  
Db 61 SAGELTLKNDLSIAALPLSCFNLGSLFTVLGRGHSITPENIRTSNGAALSADGL 120

Qy 121 FTIEGPKELSFNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180  
Db 121 FTIEGPKELSFNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180

Qy 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTADGGACQVTFSPSAMANEAPAFVAN 240  
Db 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTADGGACQVTFSPSAMANEAPAFVAN 240

Qy 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAPL 300  
Db 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAPL 300

Qy 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTNNGYDGGAIKCKGCAQ-AGSNNSGSVFD 359  
Db 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTNNGYDGGAIKCKGCAQ-AGSNNSGSVFD 359

Qy 360 GEGVVFSSNVAAGKGGAIYAKKLSVANCGPVQFLNITANDGGAIIYLGESGELSADYG 419  
Db 360 GEGVVFSSNVAAGKGGAIYAKKLSVANCGPVQFLNITANDGGAIIYLGESGELSADYG 419

Qy 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKIITLRAKAGHQLFPNDPIEMANGN 479  
Db 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKIITLRAKAGHQLFPNDPIEMANGN 480

Qy 480 NQPAQSSKLLKINDGEGVTGDI VFANGSSTLYQNTYI EQGRIVLRKAKLSVNSISQGG 539  
Db 480 NQPAQSSKLLKINDGEGVTGDI VFANGSSTLYQNTYI EQGRIVLRKAKLSVNSISQGG 540

Qy 540 SLYMEAGSTWDFVTPPQQPQPAANQLITLNLHLSSLLANNAVTPNPPNPAQDSHP 599  
Db 540 SLYMEAGSTWDFVTPPQQPQPAANQLITLNLHLSSLLANNAVTPNPPNPAQDSHP 600

Qy 600 AVIGSTTAGSVTIISGPIFEDLDYADRYDNLGSKINVLKQLGTYKPPANAPSDTL 659  
Db 600 AVIGSTTAGSVTIISGPIFEDLDYADRYDNLGSKINVLKQLGTYKPPANAPSDTL 660

Qy 660 GNEAPKYGQSGWKALWDPNTANNPPTLKATWTKTYNCPGPVRSIVLPNLSGSLDI 719  
Db 660 GNEAPKYGQSGWKALWDPNTANNPPTLKATWTKTYNCPGPVRSIVLPNLSGSLDI 720

Qy 720 RSAHSIAIQASVDGRSICRGLWVSGVSNFFYHDDRDLGQGYRIYISGYSLGANSYFGSSMF 779  
Db 720 RSAHSIAIQASVDGRSICRGLWVSGVSNFFYHDDRDLGQGYRIYISGYSLGANSYFGSSMF 780

Qy 780 GLATEVFGRSKDYVVCESNHHACIGSVYLSTQOALCGSYLFGDAFIRASVFGNQHMKT 839  
Db 780 GLATEVFGRSKDYVVCESNHHACIGSVYLSTQOALCGSYLFGDAFIRASVFGNQHMKT 840

Qy 840 SYTFAESDVRWDNNCLAGELGAGLPVITPSKLYLNELRPPVQAEPSYADHESPTESGD 899  
Db 840 SYTFAESDVRWDNNCLAGELGAGLPVITPSKLYLNELRPPVQAEPSYADHESPTESGD 900

Qy 900 QARAPKSGHLLNLSPVGVKDFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQBTW 959  
Db 900 QARAPKSGHLLNLSPVGVKDFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQBTW 960

Qy 960 TTDAPHLARHGVVVRGSMYASLTNSIEVYHGRYEDRDSRCYGLSAGSRVP 1012  
Db 960 TTDAPHLARHGVVVRGSMYASLTNSIEVYHGRYEDRDSRCYGLSAGSRVP 1013

RESULT 21  
US-10-931-779-15  
; Sequence 15, Application US/10931779  
; Publication No. US20050048557A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: BPL04  
; CURRENT APPLICATION NUMBER: US/10/931,779  
; CURRENT FILING DATE: 2004-09-01  
; PRIOR APPLICATION NUMBER: 09/542,520  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-931-779-15

Query Match 97.0%; Score 5106.5; DB 5; Length 1013;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;

Qy 1 MQTSPHKFPLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFTYVIGDPSGTTVF 60  
Db 1 MQTSPHKFPLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFTYVIGDPSGTTVF 60

Qy 61 SAGELTLKNDLSIAALPLSCFNLGSLFTVLGRGHSITPENIRTSNGAALSADGL 120  
Db 61 SAGELTLKNDLSIAALPLSCFNLGSLFTVLGRGHSITPENIRTSNGAALSADGL 120

Qy 121 FTIEGPKELSFNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180  
Db 121 FTIEGPKELSFNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180

Qy 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTADGGACQVTFSPSAMANEAPAFVAN 240  
Db 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTADGGACQVTFSPSAMANEAPAFVAN 240

Qy 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAPL 300  
Db 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAPL 300

Qy 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTNNGYDGGAIKCKGCAQ-AGSNNSGSVFD 359  
Db 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTNNGYDGGAIKCKGCAQ-AGSNNSGSVFD 360

Qy 360 GEGVVFSSNVAAGKGGAIYAKKLSVANCGPVQFLNITANDGGAIIYLGESGELSADYG 419  
Db 360 GEGVVFSSNVAAGKGGAIYAKKLSVANCGPVQFLNITANDGGAIIYLGESGELSADYG 420

Qy 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKIITLRAKAGHQLFPNDPIEMANGN 479  
Db 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKIITLRAKAGHQLFPNDPIEMANGN 480



```
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-190

Query Match      96.6%; Score 5090; DB 3; Length 1006;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 999; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 31 MYPQGIYDGETLTVSPPTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 90
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
25 MIPQGIYDGETLTVSPPTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 84
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 91 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGFKLSFNSCNLSLAVLPAATTNKS 150
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
85 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGFKLSFNSCNLSLAVLPAATTNKS 144
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 151 QTPPTTSPNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVPOE 210
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
145 QTPPTTSPNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVPOE 204
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 211 NTAQADGGACQVVTSPSAMANEAPIAFVANVAGVGGGIAAVQDQCGQVSSSTSTEDPVV 270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 NTAQADGGACQVVTSPSAMANEAPIAFVANVAGVGGGIAAVQDQCGQVSSSTSTEDPVV 264
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 271 SPSRNTAVFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNT 330
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
265 SPSRNTAVFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNT 324
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 331 SNNYDGGAIKCKNGAQAGSNNSGSVFDGEGVFPSSNVAAGKGAIYAKKLSVANCGP 390
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
325 SNNYDGGAIKCKNGAQAGSNNSGSVFDGEGVFPSSNVAAGKGAIYAKKLSVANCGP 384
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 391 VOFLRNANDGGAIYLGSGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSQAISM 450
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
385 VOFLRNANDGGAIYLGSGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSQAISM 444
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 451 GSGGKITTLRAKAGHQLFNDPIEMANGNPPAQSSKLLKINDGEGYTGDIVFANGSSTL 510
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 GSGGKITTLRAKAGHQLFNDPIEMANGNPPAQSSKLLKINDGEGYTGDIVFANGSSTL 504
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 511 YONTIEQGRIVLREKAKLSVNSLSQCGSLYMEAGSTWDFVTPQPPQPPAAQLITLS 570
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 YONTIEQGRIVLREKAKLSVNSLSQCGSLYMEAGSTLDFVTPQPPQPPAAQLITLS 564
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 571 NLHLSLLANNAVTPNPPAQDQSHPAVIGSTTAGSVTISGPIPPEDLDYDTRYD 630
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
565 NLHLSLLANNAVTPNPPAQDQSHPAVIGSTTAGSVTISGPIPPEDLDYDTRYD 624
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 631 WLGNQKINVLKQLGTGKPPANAPSDLTGNEPKYIGQSWKLAWDPTANNPPTLKA 690
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
625 WLGNQKINVLKQLGTGKPPANAPSDLTGNEPKYIGQSWKLAWDPTANNPPTLKA 684
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 691 TWTGTGNGPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLWVSGVSNFPYH 750
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
685 TWTGTGNGPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLWVSGVSNFPYH 744
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 751 DRDALGQGYRYISGGYISLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHACIGSVYLS 810
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
745 DRDALGQGYRYISGGYISLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHACIGSVYLS 804
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 811 TQOALCGSYLPDGAIFIRASYFGNQHMKTSTYTFABESDVRDNNCLAGEIGAGLPVITP 870
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
805 TQOALCGSYLPDGAIFIRASYFGNQHMKTSTYTFABESDVRDNNCLAGEIGAGLPVITP 864
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 871 SKLYLNEARLPVQAEFSAHDESFTBEGDQARAFKSGHLLNLSVPVGVKFDRCSTHPNK 930
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
865 SKLYLNEARLPVQAEFSAHDESFTBEGDQARAFKSGHLLNLSVPVGVKFDRCSTHPNK 924
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 931 YSFMAAVICDAYRTISGTETTLTLLSHQETWTTDAPHLARHGVTVRGSMYASLTNSIEVYGH 990
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Db	385	VQFLNIANDGGAIYLGSGELSLADYDGIIFDGNLAKRTAKENAAADVNTVSSQAISM	444
Qy	451	CGSGKITTLRAKAGHOIILFNDPIEMANGNPOAQSSKLLKINDGEGYTGDIIVFANGSSTL	510
Db	445	CGSGKITTLRAKAGHOIILFNDPIEMANGNPOAQSSKLLKINDGEGYTGDIIVFANGSSTL	504
Qy	511	YQNVITIEGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPPOQPPAANQILITLS	570
Db	505	YQNVITIEGRIVLREKAKLSVNSLSQSGSLYMEAGSTLDFVTPPOQPPAANQILITLS	564
Qy	571	NLHLSLSLLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD	630
Db	565	NLHLSLSLLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD	624
Qy	631	WLGSNQKINVLKQLGTRKPPANAPSDLTILGNEMPKYGYQGSWKLAWDPNTANNPPTLKA	690
Db	625	WLGSNQKINVLKQLGTRKPPANAPSDLTILGNEMPKYGYQGSWKLAWDPNTANNPPTLKA	684
Qy	691	TWTKTYGPNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	750
Db	685	TWTKTYGPNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	744
Qy	751	DRDALGQGYRYISGGYSLGANSYFGSSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS	810
Db	745	DRDALGQGYRYISGGYSLGANSYFGSSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS	804
Qy	811	TQOALCGSYLFGDAFIRASYGFGNQHMTSYTFABESDVRWNNCLAGIBIGAGLPVITP	870
Db	805	TQOALCGSYLFGDAFIRASYGFGNQHMTSYTFABESDVRWNNCLAGIBIGAGLPVITP	864
Qy	871	SKLYLNLAPFVQAEBSYADHESFTEEGQARAFKSGHLLNLSPVGVKFDRCSSSTHPNK	930
Db	865	SKLYLNLAPFVQAEBSYADHESFTEEGQARAFKSGHLLNLSPVGVKFDRCSSSTHPNK	924
Qy	931	YSFMAAICDAVRTISGTETTLSSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	990
Db	925	YSFMAAICDAVRTISGTETTLSSHQETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	984
Qy	991	GRYEYRDASRGYGLSAGSRVRF 1012	
Db	985	GRYEYRDASRGYGLSAGSRVRF 1006	
RESULT 25			
US-09-841-132-176			
; Sequence 176, Application US/09841132			
; Patent No. US20020061848A1			
; GENERAL INFORMATION:			
; APPLICANT: Bhatia, Ajay			
; APPLICANT: Skeiky, Yasir A.W.			
; APPLICANT: Probst, Peter			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND			
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION			
; FILE REFERENCE: 210121.469C8			
; CURRENT APPLICATION NUMBER: US/09/841,132			
; CURRENT FILING DATE: 2001-04-23			
; NUMBER OF SEQ ID NOS: 599			
; SOFTWARE: FastSeq for Windows Version 3.0/4.0			
; SEQ ID NO 176			
; LENGTH: 982			
; TYPE: PRT			
; ORGANISM: Chlamydia			
; FEATURE:			
; NAME/KEY: VARIANT			
; LOCATION: (1)...(982)			
; OTHER INFORMATION: Xaa = Any Amino Acid			
US-09-841-132-176			
Query Match 96.5%; Score 5084; DB 3; Length 982;			
Best Local Similarity 99.6%; Pred. No. 0;			
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;			

RESULT 26  
US-10-872-155-176  
; Sequence 176, Application US/10872155  
; Publication No. US20040234536A1





208 POENTAQADGGACQVVTTSFSAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTED 267  
181 POENTAQADGGACQVVTTSFSAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTED 240  
268 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAKQPTSGQA 327  
241 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAEQPTNGQA 300  
328 SNTSNNGYDGGAI FCKNGAQ-AGSNNSGVSFDGEGVVPFSSNVAAGKGGAIYAKKLSVA 386  
301 SNTSDNYDGGAI FCKNGAQAGSNNSGVSFDGEGVVPFSSNVAAGKGGAIYAKKLSVA 360  
387 NCGPVQFLNANDGGAIIYLGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 446  
361 NCGPVQFLNANDGGAIIYLGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 420  
447 AISMGSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANG 506  
421 AISMGSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANG 480  
507 SSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPPOPPQPPAANQL 566  
481 NSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPPOPPQPPAANQL 540  
567 ITLSNLHLSSLLANNVTPNPPAQDSSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626  
541 ITLSNLHLSSLLANNVTPNPPAQDSSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 600  
627 DRYDWMGSKNIDVLKQLGTQPSANAPSDLTILGNEMPKYGYQGSWKLAWDNTANNPGY 686  
601 DRYDWMGSKNIDVLKQLGTQPSANAPSDLTILGNEMPKYGYQGSWKLAWDNTANNPGY 660  
687 TLKATWTKTG 696  
661 TLKATWTKTG 670

RESULT 28  
US-11-109-468-169  
; Sequence 169, Application US/11109468  
; Publication No. US20050232941A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Gunderian, Jeff  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Maisonneuve, Jean-Francois L.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT  
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.515C4  
; CURRENT FILING DATE: 2005-04-19  
; PRIOR APPLICATION NUMBER: US 10/197,220  
; PRIOR FILING DATE: 2002-07-15  
; PRIOR APPLICATION NUMBER: US 10/007,693  
; PRIOR FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: US 10/012,256  
; PRIOR FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 09/841,260  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/219,752  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: US 60/198,853  
; PRIOR FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 175  
; SEQ ID NO 169  
; LENGTH: 670  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-11-109-468-169

Query Match 63.3%; Score 3336.5; DB 6; Length 670;  
Best Local Similarity 97.2%; Pred. No. 3.6e-238;

Matches 651; Conservative 11; Mismatches 7; Indels 1; Gaps 1;  
QY 28 AEIMVPOGIYDGETLTVSFPYTVIGDPSGTTVFAGELTTLKNLDSIAALPLSCFGNLLG 87  
DB 1 AEIMVPOGIYDGETLTVSFPYTVIGDPSGTTVFAGELTTLKNLDSIAALPLSCFGNLLG 60  
QY 88 SFTVLGRGHSITFENIRITSTNGAALSADGLFTIEGFKELSPSNCNLSLAVLPAATN 147  
DB 61 SFTVLGRGHSITFENIRITSTNGAALSADGLFTIEGFKELSPSNCNLSLAVLPAATN 120  
QY 148 KGSQPTTSTPSNGTIYSKTDLLILANNEKSPFYNLSVGGDGAIDAKSLTVQGISKLCV 207  
DB 121 NGSQPTTSTPSNGTIYSKTDLLILANNEKSPFYNLSVGGDGAIDAKSLTVQGISKLCV 180  
QY 208 FOENTAQADGGACQVVTTSFSAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTED 267  
DB 181 FOENTAQADGGACQVVTTSFSAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTED 240  
QY 268 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAKQPTSGQA 327  
DB 241 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAEQPTNGQA 300  
QY 328 SNTSNNGYDGGAI FCKNGAQ-AGSNNSGVSFDGEGVVPFSSNVAAGKGGAIYAKKLSVA 386  
DB 301 SNTSDNYDGGAI FCKNGAQAGSNNSGVSFDGEGVVPFSSNVAAGKGGAIYAKKLSVA 360  
QY 387 NCGPVQFLNANDGGAIIYLGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 446  
DB 361 NCGPVQFLNANDGGAIIYLGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 420  
QY 447 AISMGSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANG 506  
DB 421 AISMGSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANG 480  
QY 507 SSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPPOPPQPPAANQL 566  
DB 481 NSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPPOPPQPPAANQL 540  
QY 567 ITLSNLHLSSLLANNVTPNPPAQDSSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626  
DB 541 ITLSNLHLSSLLANNVTPNPPAQDSSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 600  
QY 627 DRYDWMGSKNIDVLKQLGTQPSANAPSDLTILGNEMPKYGYQGSWKLAWDNTANNPGY 686  
DB 601 DRYDWMGSKNIDVLKQLGTQPSANAPSDLTILGNEMPKYGYQGSWKLAWDNTANNPGY 660  
QY 687 TLKATWTKTG 696  
DB 661 TLKATWTKTG 670

RESULT 29  
US-10-701-844-17  
; Sequence 17, Application US/10701844  
; Publication No. US20040067524A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, W. James  
; APPLICANT: Pace, John  
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
; FILE REFERENCE: 7969-086-999  
; CURRENT APPLICATION NUMBER: US/10/701,844  
; CURRENT FILING DATE: 2003-11-04  
; PRIOR APPLICATION NUMBER: US/09/612,402B  
; PRIOR FILING DATE: 2000-07-06  
; PRIOR APPLICATION NUMBER: 08/942,596  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Chlamydia sp.  
US-10-701-844-17

```
Query Match      48.5%; Score 2552; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 3.3e-180;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPGIYDGETLTVSPFTYVIGDPGTTVFSAGELTKNLDNSIAALPLSCFNLGSS 88
    |||
DB 1 EIMVPGIYDGETLTVSPFTYVIGDPGTTVFSAGELTKNLDNSIAALPLSCFNLGSS 60

QY 89 FTVLGRGHSITPENIRTSNGAALSNSAADGLFTIEGFKELSFNCNLSLLAVLPAATTNK 148
    |||
DB 61 FTVLGRGHSITPENIRTSNGAALSNSAADGLFTIEGFKELSFNCNLSLLAVLPAATTNK 120

QY 149 GSQPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVP 208
    |||
DB 121 GSQPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVP 180

QY 269 VVSFRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 328
    |||
DB 241 VVSFRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300

QY 329 NTSNNYGDGGAIFCKNGAQAAGNNSGSVSPDGEVVFSSNVAAAGKGGAIIYAKKLSVANC 388
    |||
DB 301 NTSNNYGDGGAIFCKNGAQAAGNNSGSVSPDGEVVFSSNVAAAGKGGAIIYAKKLSVANC 360

QY 389 GPVQPLRNANDGGAIIYLGESGELSLSADYDGIIFDGNLKRKTAKENAADVNGVTSSQAI 448
    |||
DB 361 GPVQPLRNANDGGAIIYLGESGELSLSADYDGIIFDGNLKRKTAKENAADVNGVTSSQAI 420

QY 449 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 508
    |||
DB 421 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 480

QY 509 TLYQNVTIEQGRIVLREKAKLSVNS 533
    |||
DB 481 TLYQNVTIEQGRIVLREKAKLSVNS 505

RESULT 30
US-10-766-711-17
; Sequence 17, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-17

Query Match      48.5%; Score 2552; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 3.3e-180;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPGIYDGETLTVSPFTYVIGDPGTTVFSAGELTKNLDNSIAALPLSCFNLGSS 88
    |||
DB 1 EIMVPGIYDGETLTVSPFTYVIGDPGTTVFSAGELTKNLDNSIAALPLSCFNLGSS 60

QY 89 FTVLGRGHSITPENIRTSNGAALSNSAADGLFTIEGFKELSFNCNLSLLAVLPAATTNK 148
    |||
DB 61 FTVLGRGHSITPENIRTSNGAALSNSAADGLFTIEGFKELSFNCNLSLLAVLPAATTNK 120

QY 149 GSQPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVP 208
    |||
DB 121 GSQPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVP 180

QY 269 VVSFRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 328
    |||
DB 241 VVSFRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300

QY 329 NTSNNYGDGGAIFCKNGAQAAGNNSGSVSPDGEVVFSSNVAAAGKGGAIIYAKKLSVANC 388
    |||
DB 301 NTSNNYGDGGAIFCKNGAQAAGNNSGSVSPDGEVVFSSNVAAAGKGGAIIYAKKLSVANC 360

QY 389 GPVQPLRNANDGGAIIYLGESGELSLSADYDGIIFDGNLKRKTAKENAADVNGVTSSQAI 448
    |||
DB 361 GPVQPLRNANDGGAIIYLGESGELSLSADYDGIIFDGNLKRKTAKENAADVNGVTSSQAI 420

QY 449 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 508
    |||
DB 421 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 480

QY 509 TLYQNVTIEQGRIVLREKAKLSVNS 533
    |||
DB 481 TLYQNVTIEQGRIVLREKAKLSVNS 505

RESULT 31
US-10-931-779-17
; Sequence 17, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: Bp104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-17

Query Match      48.5%; Score 2552; DB 5; Length 505;
Best Local Similarity 100.0%; Pred. No. 3.3e-180;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPGIYDGETLTVSPFTYVIGDPGTTVFSAGELTKNLDNSIAALPLSCFNLGSS 88
    |||
DB 1 EIMVPGIYDGETLTVSPFTYVIGDPGTTVFSAGELTKNLDNSIAALPLSCFNLGSS 60

QY 89 FTVLGRGHSITPENIRTSNGAALSNSAADGLFTIEGFKELSFNCNLSLLAVLPAATTNK 148
    |||
DB 61 FTVLGRGHSITPENIRTSNGAALSNSAADGLFTIEGFKELSFNCNLSLLAVLPAATTNK 120

QY 149 GSQPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVP 208
    |||
DB 121 GSQPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVP 180

QY 209 QENTAQAADGGACQVVTTSFSAANEAPAFVANVAGVGGGIAAVQDQGGQVSSSTSTEDP 268
    |||
DB 181 QENTAQAADGGACQVVTTSFSAANEAPAFVANVAGVGGGIAAVQDQGGQVSSSTSTEDP 240

QY 269 VVSFRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 328
    |||
DB 241 VVSFRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300

QY 329 NTSNNYGDGGAIFCKNGAQAAGNNSGSVSPDGEVVFSSNVAAAGKGGAIIYAKKLSVANC 388
    |||
DB 301 NTSNNYGDGGAIFCKNGAQAAGNNSGSVSPDGEVVFSSNVAAAGKGGAIIYAKKLSVANC 360

QY 389 GPVQPLRNANDGGAIIYLGESGELSLSADYDGIIFDGNLKRKTAKENAADVNGVTSSQAI 448
    |||
DB 361 GPVQPLRNANDGGAIIYLGESGELSLSADYDGIIFDGNLKRKTAKENAADVNGVTSSQAI 420

QY 449 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 508
    |||
DB 421 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 480

QY 509 TLYQNVTIEQGRIVLREKAKLSVNS 533
    |||
DB 481 TLYQNVTIEQGRIVLREKAKLSVNS 505
```

Db 361 SSSLANNNAVTPPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 420

Qy 637 KINVLKQLQGTGKTPANAPSDLTGLNEMPKYGYQGSWKL 674

Db 421 KINVLKQLQGTGKTPANAPSDLTGLNEMPKYGYQGSWKL 458

RESULT 33

US-10-766-711-36

Sequence 36, Application US/10766711

Publication No. US20040137005A1

GENERAL INFORMATION:

APPLICANT: Jackson, W. James

APPLICANT: Pace, John

TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof

FILE REFERENCE: 7969-062-999

CURRENT APPLICATION NUMBER: US/10/766,711

CURRENT FILING DATE: 2004-01-27

PRIOR APPLICATION NUMBER: US/08/942,596

PRIOR FILING DATE: 1997-10-02

NUMBER OF SEQ ID NOS: 43

SOFTWARE: PatentIn version 3.0

SEQ ID NO 36

LENGTH: 458

TYPE: PRT

ORGANISM: Chlamydia sp.

US-10-766-711-36

Query Match 44.6%; Score 2350; DB 4; Length 458;

Best Local Similarity 100.0%; Pred. No. 2.7e-165;

Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 217 GCACQVVTSPSANEAPAFVANVAGVGGGIAAVQDGGQVSSSTSTEDPVSFSRNT 276

Db 1 GCACQVVTSPSANEAPAFVANVAGVGGGIAAVQDGGQVSSSTSTEDPVSFSRNT 60

Qy 277 AVEFDGNVARVGGGIYSGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 336

Db 61 AVEFDGNVARVGGGIYSGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 120

Qy 337 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLRN 396

Db 121 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLRN 180

Qy 397 IANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSSQAISMGSQGI 456

Db 181 IANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSSQAISMGSQGI 240

Qy 457 TTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTI 516

Db 241 TTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTI 300

Qy 517 EQGRIVLRKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHSL 576

Db 301 EQGRIVLRKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHSL 360

Qy 577 SSSLANNNAVTPPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 636

Db 361 SSSLANNNAVTPPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 420

Qy 637 KINVLKQLQGTGKTPANAPSDLTGLNEMPKYGYQGSWKL 674

Db 421 KINVLKQLQGTGKTPANAPSDLTGLNEMPKYGYQGSWKL 458

RESULT 34

US-10-931-779-36

Sequence 36, Application US/10931779

Publication No. US20050048557A1

GENERAL INFORMATION:

APPLICANT: Jackson, W. James

APPLICANT: Pace, John

TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof

Db 241 VVSFERTAVFDGNVARVGGGIYSGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300

Qy 329 NTSNNYGDGGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIAKLSVANC 388

Db 301 NTSNNYGDGGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIAKLSVANC 360

Qy 389 GPVQFLRNTANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI 448

Db 361 GPVQFLRNTANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI 420

Qy 449 SMGSGGKITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGS 508

Db 421 SMGSGGKITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGS 480

Qy 509 TLYQNVTIEQGRIVLRKAKLSVNS 533

Db 481 TLYQNVTIEQGRIVLRKAKLSVNS 505

RESULT 32

US-10-701-844-36

Sequence 36, Application US/10701844

Publication No. US20040067524A1

GENERAL INFORMATION:

APPLICANT: Jackson, W. James

APPLICANT: Pace, John

TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof

FILE REFERENCE: 7969-086-999

CURRENT APPLICATION NUMBER: US/10/701,844

CURRENT FILING DATE: 2003-11-04

PRIOR APPLICATION NUMBER: US/09/612,402B

PRIOR FILING DATE: 2000-07-06

PRIOR APPLICATION NUMBER: 08/942,596

PRIOR FILING DATE: 1997-10-02

NUMBER OF SEQ ID NOS: 43

SOFTWARE: PatentIn version 3.0

SEQ ID NO 36

LENGTH: 458

TYPE: PRT

ORGANISM: Chlamydia sp.

US-10-701-844-36

Query Match 44.6%; Score 2350; DB 4; Length 458;

Best Local Similarity 100.0%; Pred. No. 2.7e-165;

Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 217 GCACQVVTSPSANEAPAFVANVAGVGGGIAAVQDGGQVSSSTSTEDPVSFSRNT 276

Db 1 GCACQVVTSPSANEAPAFVANVAGVGGGIAAVQDGGQVSSSTSTEDPVSFSRNT 60

Qy 277 AVEFDGNVARVGGGIYSGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 336

Db 61 AVEFDGNVARVGGGIYSGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 120

Qy 337 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLRN 396

Db 121 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLRN 180

Qy 397 IANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSSQAISMGSQGI 456

Db 181 IANDGGAIIYLGESGELSADYDGIIFDGNLKRKTAKENAADVNGVTVSSQAISMGSQGI 240

Qy 457 TTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTI 516

Db 241 TTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTI 300

Qy 517 EQGRIVLRKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHSL 576

Db 301 EQGRIVLRKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHSL 360

Qy 577 SSSLANNNAVTPPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 636

```
/ FILE REFERENCE: BP104
/ CURRENT APPLICATION NUMBER: US/10/931,779
/ CURRENT FILING DATE: 2004-09-01
/ PRIOR APPLICATION NUMBER: 09/542,520
/ PRIOR FILING DATE: 200-04-03
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 36
/ LENGTH: 458
/ TYPE: PRT
/ ORGANISM: Chlamydia sp.
US-10-931-779-36

Query Match      44.6%; Score 2350; DB 5; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.7e-185;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GGACQVVTSPFSAMANEAFIAFVANVAGVGGGIAAQQDQGVSSSTSTEDPVVPSFRNT 276
DB 1 GGACQVVTSPFSAMANEAFIAFVANVAGVGGGIAAQQDQGVSSSTSTEDPVVPSFRNT 60

QY 277 AVFPGNVARVGGGYSYGNVAFLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 336
DB 61 AVFPGNVARVGGGYSYGNVAFLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 120

QY 337 GGAIFCKGAQAGSNNSVSFDEGVVFFPSNVAAGGAIYAKLSVANCGPVQFLRN 396
DB 121 GGAIFCKGAQAGSNNSVSFDEGVVFFPSNVAAGGAIYAKLSVANCGPVQFLRN 180

QY 397 IANDGAIYLGESGLSADYDGIIFDGNLKRTAKENAADVNGVTVSQAISMGGGKI 456
DB 181 IANDGAIYLGESGLSADYDGIIFDGNLKRTAKENAADVNGVTVSQAISMGGGKI 240

QY 457 TTLAKAGHQLFNDPIEMANGNPPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTI 516
DB 241 TTLAKAGHQLFNDPIEMANGNPPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTI 300

QY 517 EGRIVLREKAKLSVNSLSOTGGSLYMEAGSTWDFVTPQPPAANQLITLSNLHLSL 576
DB 301 EGRIVLREKAKLSVNSLSOTGGSLYMEAGSTWDFVTPQPPAANQLITLSNLHLSL 360

QY 577 SLLANNAVTPPTNPQAQDHPAVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNQ 636
DB 361 SLLANNAVTPPTNPQAQDHPAVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNQ 420

QY 637 KINVLKQLGKTPKANAPSDLTGLNEMPKYGYQGSWKL 674
DB 421 KINVLKQLGKTPKANAPSDLTGLNEMPKYGYQGSWKL 458

RESULT 35
US-10-701-844-37
/ Sequence 37, Application US/10701844
/ Publication No. US20040067524A1
/ GENERAL INFORMATION:
/ APPLICANT: Jackson, W. James
/ TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
/ FILE REFERENCE: 7969-086-999
/ CURRENT APPLICATION NUMBER: US/10/701,844
/ CURRENT FILING DATE: 2003-11-04
/ PRIOR APPLICATION NUMBER: US/09/612,402B
/ PRIOR FILING DATE: 2000-07-06
/ PRIOR APPLICATION NUMBER: 08/942,596
/ PRIOR FILING DATE: 1997-10-02
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 37
/ LENGTH: 325
/ TYPE: PRT
/ ORGANISM: Chlamydia sp.
US-10-701-844-37

Query Match      32.9%; Score 1735; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.2e-120;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKTGNPGRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 747
DB 1 LKATWTKTGNPGRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 60

QY 748 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNNHACIGSV 807
DB 61 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNNHACIGSV 120

QY 808 YLSTQQAALCGSYLPGDAFIRASVYFGNQHMKTSYTFPABESDVRDNNCLAGEICAGLPV 867
DB 121 YLSTQQAALCGSYLPGDAFIRASVYFGNQHMKTSYTFPABESDVRDNNCLAGEICAGLPV 180

QY 868 ITPSKLYLNELRPPVQAEFSDHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 927
DB 181 ITPSKLYLNELRPPVQAEFSDHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 240

QY 928 PNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 987
DB 241 PNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 300

Query Match      32.9%; Score 1735; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.2e-120;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKTGNPGRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 747
DB 1 LKATWTKTGNPGRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 60

QY 748 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNNHACIGSV 807
DB 61 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNNHACIGSV 120

QY 808 YLSTQQAALCGSYLPGDAFIRASVYFGNQHMKTSYTFPABESDVRDNNCLAGEICAGLPV 867
DB 121 YLSTQQAALCGSYLPGDAFIRASVYFGNQHMKTSYTFPABESDVRDNNCLAGEICAGLPV 180

QY 868 ITPSKLYLNELRPPVQAEFSDHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 927
DB 181 ITPSKLYLNELRPPVQAEFSDHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 240

QY 928 PNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 987
DB 241 PNKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 300
```

Thu May 25 10:14:23 2006

1012 YHGRYEYRDASRGYGLSAGSRVP 1012  
301 YHGRYEYRDASRGYGLSAGSRVP 325  
RESULT 37  
US-10-931-779-37  
Sequence 37, Application US/10931779  
Publication No. US20050048557A1  
GENERAL INFORMATION:  
APPLICANT: Pace, John  
TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
FILE REFERENCE: BP104  
CURRENT APPLICATION NUMBER: US/10/931,779  
PRIORITY FILING DATE: 2004-09-01  
PRIORITY FILING DATE: 09/542,520  
PRIORITY FILING DATE: 200-04-03  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 37  
LENGTH: 325  
TYPE: PRT  
ORGANISM: Chlamydia sp.  
US-10-931-779-37  
Query Match 32.9%; Score 1735; DB 5; Length 325;  
Best Local Similarity 100.0%; Pred. No. 6.2e-120;  
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 688 LKATWTKTGNPCGPVAVSLVPSNLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSVNF 747  
DB 1 LKATWTKTGNPCGPVAVSLVPSNLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSVNF 60  
QY 748 FYHDDRALGGYRIYISGGYSLGANSYFGSSMFLAFTVEFGRSKDYVVCRSNHHACIGSV 807  
DB 61 FYHDDRALGGYRIYISGGYSLGANSYFGSSMFLAFTVEFGRSKDYVVCRSNHHACIGSV 120  
QY 808 YLSTQALCGSYLFGDAFIRASYGFGNQHMKTSTYTPAESDVRWDDNCLAGEIGAGLPV 867  
DB 121 YLSTQALCGSYLFGDAFIRASYGFGNQHMKTSTYTPAESDVRWDDNCLAGEIGAGLPV 180  
QY 868 IPTSKLYNELRPFVQAEFSYADHESFTEEGDQARAFKSGHLLNLSPVGVKFDRCSSSTH 927  
DB 181 IPTSKLYNELRPFVQAEFSYADHESFTEEGDQARAFKSGHLLNLSPVGVKFDRCSSSTH 240  
QY 928 PNKYSFMAAYICDAYRTISGTETLLSHQETWTTDAPHLARHGTVVGRGSMYASLTGNIEV 987  
DB 241 PNKYSFMAAYICDAYRTISGTETLLSHQETWTTDAPHLARHGTVVGRGSMYASLTGNIEV 300  
QY 988 YHGRYEYRDASRGYGLSAGSRVP 1012  
DB 301 YHGRYEYRDASRGYGLSAGSRVP 325  
RESULT 38  
US-10-312-273-3  
Sequence 38, Application US/10312273  
Publication No. US20040005667A1  
GENERAL INFORMATION:  
APPLICANT: CHIRON SpA  
TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE  
FILE REFERENCE: P025035W0  
CURRENT APPLICATION NUMBER: US/10/312,273  
PRIORITY FILING DATE: 2002-12-20  
PRIORITY FILING DATE: 0016363.4  
PRIORITY FILING DATE: 2000-07-03  
PRIORITY FILING DATE: 0017047.2  
PRIORITY FILING DATE: 2000-07-11  
PRIORITY FILING DATE: 0017983.8  
PRIORITY FILING DATE: 2000-07-21  
PRIORITY FILING DATE: 0019368.0

1012 YHGRYEYRDASRGYGLSAGSRVP 1012  
301 YHGRYEYRDASRGYGLSAGSRVP 325  
PRIORITY FILING DATE: 2000-08-07  
PRIORITY APPLICATION NUMBER: 0020440.4  
PRIORITY FILING DATE: 2000-08-18  
PRIORITY APPLICATION NUMBER: 0022583.9  
PRIORITY FILING DATE: 2000-09-14  
PRIORITY APPLICATION NUMBER: 0027549.5  
PRIORITY FILING DATE: 2000-11-10  
PRIORITY APPLICATION NUMBER: 0031706.5  
PRIORITY FILING DATE: 2000-12-22  
NUMBER OF SEQ ID NOS: 664  
SOFTWARE: SeqWin99, version 1.02  
SEQ ID NO 3  
LENGTH: 973  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-10-312-273-3  
Query Match 25.9%; Score 1365.5; DB 4; Length 973;  
Best Local Similarity 33.5%; Pred. No. 7.6e-92;  
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;  
QY 1 MOTSPHKPFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPFVTVIGDPGSGTVF 60  
DB 1 MKTSIRKFLISTTLAPCPAST---AFTVEVIMPSNFSGSOKI--FPYTTLSDDPGTLCI 56  
QY 61 SAGELTLKNDLNSIAALPLSCFENLLGSFTVLGRHSLTFENIRTSNTGAALNSAAGD- 119  
DB 57 FSGDLYIANLDNAISRTSSCFNFRAGALQILKGVFSFLNIRSSADCAAISSVITQNP 116  
QY 120 ---LFTISGPKEFLSPNCSNLLAVLPAATNKGSGOTPTTTPSNGTIYSKTDILLNNE 176  
DB 117 ELCPLSPSGFSQMI FDNCESLT-----SDTSASNVIPHASAIYATTPMLFTNND 165  
QY 177 KFSFYNLVSVDGGGDAIDAKSLTVQGISKLCVQENTAGDGCACQVTVSFSAMANEAPTA 236  
DB 166 SILFOYNSAGFGAAIRGTSTIETNTKSLLPNGNSISNGALTGSAININNSAPVI 225  
QY 237 FVANVAGVRGGIAAVQDQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVGGIYSYGN 296  
DB 226 FSTNATGIYGGAIYLT-----GNSLGVLFVNNSSRSGAIYANGN 272  
QY 297 VAFLLNNGKTLFANNVASPV-YIAAKOPTSGQASNTSNVVDGGAIFCKNGAQAQSNSSGS 355  
DB 273 VTFNNSDLTFQNTTASPNLSLPAPTPPPTPAVTP-LILGYGAIFFCTPPATPTPTGV-S 330  
QY 356 VSFDEGVVFPSSNVAAGKGAIAKLSVANGCPVQFLRNITANDGGAIYLGESBELSUS 415  
DB 331 LTISGENSVTFLENTASEQGGALYKKISIDSNKSTIFLIGNTAGKGAIAIPESBELSUS 390  
QY 416 ADYGDIIPGNLKRITAKENAAVNGVTVSSQAI SMGSGGKIITLRAKAGHQLFNDPIEM 475  
DB 391 ANQGDILFNKLSITSG-----TPRNSIHFGKDAKFATLGATQGYTLVFDYDPI-- 439  
QY 476 ANGNQPAQSSKLLKIN-----DGEYTGDIYF-----ANGSSLYQNVVTR 517  
DB 440 TSDDLASAASAAATVVVNPKASADG-AYSCTIVFSGEITLTAETAATPANATSTLNQKLE 498  
QY 518 QGRIVLRKAKLSVNSLSQTGS-LYMEAGSTWDFVTPPOPPQPPAANQLITLNLHLSL 576  
DB 499 GGTALRNGATLNVHNFOTDEKSVWIMDAGTT--LATNGANNTOGA---ITLNLKLVINL 553  
QY 577 SSLANNNAVTPNPTPPAQDSDHPAVIG-STTAGSVTISGPIFFEDLDDTAYDRYDWLGSN 635  
DB 554 DSLDGTKA-----AVNVOSTNGALITISGTGLVKNSSQCCDNHGMFNKD 598  
QY 636 -QKINVILQLQGTCKPPANAPSDL-TLGNEMPKYGYQGSWKLAWDPNTANGPYTLKATWT 693  
DB 599 LQQVPIELKATSNVTVTTFDLSLTNGYQSPYQGTWETIDTFTT-----HTVTGNMK 653  
QY 694 KTGYNPGERVASLVNSLWGSILDIRSAHSAIQASVDGRSY-CRGLWTVSGVSNFPHDR 752  
DB 654 KTGYPHPRPLAPLIPNSLWANNVIDLRAVSA--SAADGEDVPGKLSITGITNFFPHANH 711



Query Match		21.4%;	Score.1128;	DB.3;	Length.928;
Best Local Similarity		29.2%;	Pred.No.2.8e-74;		
Matches.311;		Conservative.169;	Mismatch.394;	Indels.192;	Gaps.32;
Qy	1	MOTSFKPFLSMILAYSCSLNGGGYAAE-IMVPOGIYDGETLTVSPFYTVIGDPSGTTV	59		
Db	1	MKSSPKFVFSTFAIFPLSMI-----ATETVLDSSASFDGNK-NGNFSVRESQEDAGTTY	54		
Qy	60	FSAGELTLKVLNLSIAALPLSCFNLGSLFTVLGRGHSITFENIRTS-NGAALNSAAD	118		
Db	55	LFGKNTVLENIPTGTAITKSCFNNTKGLTFTGNGNSLLFQTVDA GTVAGAAVNSSVVD	114		
Qy	119	GUFTIEGPKELFSNCNSLLAVLPAATTNKGSGTPTTTTPSPNGTIYSKTDLLLLNNEKF	178		
Db	115	KSTTFPGSLSF-----IASPGSSITTKGAVSCSTGS-----LSLTKNVL	157		
Qy	179	SPYSNLVSGDGAIDAKSLTVQISKLCVFQENTAQDGGACQVTFPSAMANEAPIAFV	238		
Db	158	LFSKNFSTDNGGAI TAKTILSLGTWMSALFSENSSKKGAIQTS DALTITGNQGEVSF-	216		
Qy	239	ANVAGVRGGGIAAVDQGGQSSSTSTEDPVVSFSRNTAVEPDGNVARVGGGIYSYGNVA	298		
Db	217	-----SDNTSSDS-----GAAIFTEASVT	235		
Qy	299	FLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNYDGGAI FCKNGAQAGSNNSGSVSF	358		
Db	236	ISNAKVSEIDN-----KVTGASSSTGDM-SGGAI CAYK-----TSTDTKVTL	278		
Qy	359	DGEGVVPFSSNAAGKGGAIYAKKLSVANGCPVQFLRNIA-----DGGAIYLGESGELS	413		
Db	279	TGNQMLFNSNTTTAGGAIYVKKLELASGELTLPFRNSVNGGTAPKGGAI AIEDSGELS	338		
Qy	414	LSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAISWGGCKITTLRAXAGHOILFNDPI	473		
Db	339	LSADSGDIVPLGN---TVSTTPTCTN---RGSIDLGTSAKMTALRSAAGRAIYFYDPI	390		
Qy	474	EMANGNNQPAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNTV	515		
Db	391	TTGSSTT-----VTDVLKVNETPADSALQYTGNIIFTGEKLS ETEAADSKNLT SKLIQPV	446		
Qy	516	IEGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPPQPPQPAANQLITLGNLHL	574		
Db	447	LSGGTSLKHGVTLQOAF TQOADSRL EMDVGITL E-----PADTS--TINNVI	494		
Qy	575	SLSSLLANNAVTPPTPPAQDHPAVIGS--TTAGSVTISGPIFFEDLDLTAYDRYDWLG	633		
Db	495	NISSI-----DGAKKAKIETKATSKNLTLSGTTITLLDPTGTFYENHS-LR	538		
Qy	634	SNQKINVLKLOL-GTKPPANAPSDLTGNEMPKYQGSW-KLAWDPNTANNGPYTLKAT	691		
Db	539	NPQSYDILELKASGTVTSTAVTPDPIMGEKF-HYGYQGTWGPVIMGTGASTTATF-----N	593		
Qy	692	WTKTYNPGPERVASLPNSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFPFYHD	751		
Db	594	WTKGYIPNPERIGSLVPSLWNAFIDISSLHYLMETANEGLQGDRAFCAGLSNPFHKD	653		
Qy	752	RDALGQGVRYISGGYSLGANSYFGS-SMFGLAFTVEFGRSKDYVVCRSNHHACIGSYLS	810		
Db	654	STKTRGRFRLHSGGVYVIGGNLHTCSDKILSAAFCQLFGRDRDYFVAKNQ-----GT	708		
Qy	811	TQOALCGSYLFGDAFI-----RASYGFGNQHMKTSYTFABE	846		
Db	709	TLX-----YQHNETHYISLPCKLRPCSLSYVPTEIPVLPSGNLSYTHTDNDLKTYYTPT	763		
Qy	847	SDVFWDNCLAGEIGAGLPIVITPSKLYLNLRLPFOAEFSYADHESFTREGDQARAFKS	906		
Db	764	VKSGWGNDSFALBFEGGRAPICLDESALF-EQYMPFMKLQFVYAHQEGFKEQTEAREFGS	822		
Qy	907	GHLNLSPVGVKPDRCSTHPNKYSFMAAYICDAYRTISGTETTLLSHQETWTTDAFHL	966		
Db	823	SRLVNLALPIGIRFDKESDCQDATYNTLGTVDLVRSPDCTTTLRISGDSWKTFTGNL	882		

Qy 967 ARHGVVVRGSMYASLTSNIEVYGHGRVEYRDASRGYGLSAGSRVRF 1012  
Db 883 ARQALVLRAGNHFCFNSNFAPFSQFSPFLRGSSRNYNVDLGAKYQF 928

Search completed: May 13, 2006, 12:35:01  
Job time : 176 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:32:24 ; Search time 30 Seconds  
(without alignments)  
1583.741 Million cell updates/sec

Title: US-10-701-844-2

Perfect score: 5267

Sequence: 1 MQTSFHKPFLSMILAYSCS.....YEURDASRGYLSAGSRVRF 1012

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 250354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SID55/ptodata/2/pubpaa/US08\_NEW\_PUB.pep1.\*
- 2: /SID55/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /SID55/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /SID55/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /SID55/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /SID55/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 7: /SID55/ptodata/2/pubpaa/US09\_NEW\_PUB.pep1.\*
- 8: /SID55/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 9: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 10: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*
- 11: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*
- 12: /SID55/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5131.5	97.4	1013	US-11-103-957-9	Sequence 9, Appli
2	5131.5	97.4	1013	US-11-018-868-21	Sequence 21, Appl
3	771.5	14.6	878	US-11-103-957-7	Sequence 7, Appli
4	771.5	14.6	878	US-11-018-868-23	Sequence 23, Appl
5	649.5	12.3	964	US-11-103-957-13	Sequence 13, Appl
6	649.5	12.3	964	US-11-018-868-19	Sequence 19, Appl
7	615	11.7	1016	US-11-103-957-41	Sequence 41, Appl
8	615	11.7	1016	US-11-018-868-22	Sequence 22, Appl
9	597.5	11.3	1531	US-11-103-957-15	Sequence 15, Appl
10	597.5	11.3	1531	US-11-018-868-18	Sequence 18, Appl
11	551	10.5	975	US-11-103-957-47	Sequence 47, Appl
12	551	10.5	975	US-11-018-868-15	Sequence 15, Appl
13	551	10.5	975	US-11-018-868-50	Sequence 50, Appl
14	522.5	9.9	1034	US-11-103-957-11	Sequence 11, Appl
15	522.5	9.9	1034	US-11-018-868-20	Sequence 20, Appl
16	475.5	9.0	1751	US-11-103-957-45	Sequence 45, Appl
17	475.5	9.0	1751	US-11-018-868-16	Sequence 16, Appl
18	412	7.8	1770	US-11-103-957-21	Sequence 21, Appl
19	412	7.8	1770	US-11-018-868-17	Sequence 17, Appl
20	237.5	4.5	1268	US-11-052-554A-1	Sequence 1, Appli
21	232.5	4.4	1643	US-11-052-554A-172	Sequence 172, App

22	228	4.3	1250	11	US-11-052-554A-16	Sequence 16, Appl
23	218	4.1	1571	11	US-11-052-554A-2	Sequence 2, Appli
24	218	4.1	3132	11	US-11-087-099-1245	Sequence 1245, Ap
25	218	4.1	3194	11	US-11-052-554A-90	Sequence 90, Appl
26	213	4.0	5291	11	US-11-052-554A-281	Sequence 281, App
27	204.5	3.9	2902	11	US-11-052-554A-91	Sequence 91, Appl
28	194	3.7	1474	11	US-11-067-260-18	Sequence 18, Appl
29	193.5	3.7	2340	11	US-11-052-554A-171	Sequence 171, App
30	193	3.7	1468	9	US-10-467-657-1088	Sequence 1088, Ap
31	193	3.7	2399	11	US-11-052-554A-92	Sequence 92, Appl
32	190	3.6	955	11	US-11-052-554A-179	Sequence 179, App
33	190	3.6	1392	11	US-11-067-260-24	Sequence 24, Appl
34	190	3.6	1447	11	US-11-067-260-22	Sequence 22, Appl
35	189.5	3.6	1461	11	US-11-052-554A-283	Sequence 283, App
36	189.5	3.6	1981	11	US-11-045-208-38	Sequence 38, Appl
37	189.5	3.6	2015	11	US-11-052-554A-374	Sequence 374, App
38	188.5	3.6	1312	11	US-11-067-260-20	Sequence 20, Appl
39	187.5	3.6	949	11	US-11-052-554A-6	Sequence 6, Appli
40	186.5	3.5	1343	11	US-11-052-554A-284	Sequence 284, App
41	184.5	3.5	1574	11	US-11-212-443-179	Sequence 179, App
42	184.5	3.5	1978	11	US-11-212-443-60	Sequence 60, Appl
43	182	3.5	596	8	US-10-196-749-310	Sequence 310, App
44	182	3.5	596	9	US-10-063-703-100	Sequence 100, App
45	182	3.5	596	9	US-10-194-487-310	Sequence 310, App

#### ALIGNMENTS

#### RESULT 1

US-11-103-957-9  
; Sequence 9, Application US/11103957  
; Publication No. US20050281847A1  
; GENERAL INFORMATION:  
; APPLICANT: Berthet, Francois-Xavier Jacques  
; APPLICANT: Lobet, Yves  
; APPLICANT: Poolman, Jan  
; APPLICANT: Verlant, Vincent Georges Christian Louis  
; TITLE OF INVENTION: Vaccine Composition  
; FILE REFERENCE: B45261  
; CURRENT FILING DATE: 2005-04-12  
; PRIOR FILING DATE: US/10/467,534  
; PRIOR FILING DATE: 2004-02-03  
; PRIOR APPLICATION NUMBER: PCT/EP02/01356  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: GB 0103169.9  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-11-103-957-9

Query Match 97.4%; Score 5131.5; DB 11; Length 1013;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY	1	MQTSFHKPFLSMILAYSCSLNGGVAABIMVPOQIYDGETLTVSFPPYVIGDPSGTVTF	60
Db	1	MQTSFHKPFLSMILAYSCSLNGGVAABIMVPOQIYDGETLTVSFPPYVIGDPSGTVTF	60
QY	61	SAGELTKNLNDNSIAALPLSCFNLGSGFTVLGRGHSITFFENIRTSNKAALNSAAGSL	120
Db	61	SAGELTKNLNDNSIAALPLSCFNLGSGFTVLGRGHSITFFENIRTSNKAALNSAAGSL	120
QY	121	FTIEGKELSFNCNLSLLAVLPAATYTKGSGQPTTSTPSNGTIYSKTDLLILNNEKFSF	180
Db	121	FTIEGKELSFNCNLSLLAVLPAATYTKGSGQPTTSTPSNGTIYSKTDLLILNNEKFSF	180
QY	181	YSNLYSGDGAIDAKSLTVQGI SKLCVFOENTAQDGGACQVTVSFSAMANEAPITAFVN	240

```
|||||
181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFQENTAQDGGACQVTSFSAMANEAPAFIAN 240
QY VAGVGGGIAAIVQDQGGQVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
Db VAGVGGGIAAIVQDQGGQVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
QY NGKTLFLNNVSPVYIAAKQPTSGQASNTNNYDGGGAIIFCKNGAQ-AGNNSGVSFPD 359
Db NGKTLFLNNVSPVYIAAEQPTNGQASNTSDNYDGGGAIIFCKNGAQAGNNSGVSFPD 360
QY GEGVVPSSVAAGKGGAIIYAKKLSVANCQVPQFLNNTANDGGAIIYLGESGELSADYG 419
Db GEGVVPSSVAAGKGGAIIYAKKLSVANCQVPQFLNNTANDGGAIIYLGESGELSADYG 420
QY DIIFDGNLKRKTAKENAADVNGTVSSQAIISMGSGGKITTLRAKAGHQILFNDPIEMANGN 479
Db DIIFDGNLKRKTAKENAADVNGTVSSQAIISMGSGGKITTLRAKAGHQILFNDPIEMANGN 480
QY NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSOTGG 539
Db NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSOTGG 540
QY SLYNEAGSTWDFVTPQPPQPPAAANQLITLSNHLSSLANNNAVTPNPPAQDSHP 599
Db SLYNEAGSTWDFVTPQPPQPPAAANQLITLSNHLSSLANNNAVTPNPPAQDSHP 600
QY AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 659
Db AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 660
QY GNEMPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVVASIVPNSLWGSILDI 719
Db GNEMPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVVASIVPNSLWGSILDI 720
QY RSAHSAIOASVDGRSVCYCRGLWVSGVSNFFYHDDRDLGGQYRIISGGYSLGANSYFGSSMF 779
Db RSAHSAIOASVDGRSVCYCRGLWVSGVSNFFYHDDRDLGGQYRIISGGYSLGANSYFGSSMF 780
QY GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVGFQGNQHMT 839
Db GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVGFQGNQHMT 840
QY SYTFABESDVRWDDNCLVGEIGVGLPIVITPSSKLYLNELRPFVQAEPFSAHDESFTREGD 899
Db SYTFABESDVRWDDNCLVGEIGVGLPIVITPSSKLYLNELRPFVQAEPFSAHDESFTREGD 900
QY QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHOBTW 959
Db QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHOBTW 960
QY TTDAPFLARHGVIIRGSMYASLTSNIEVYGHGREYRDSAGYGLSAGSRVRF 1012
Db TTDAPFLARHGVIIRGSMYASLTSNIEVYGHGREYRDSAGYGLSAGSRVRF 1013
```

## RESULT 2

```
US-11-018-868-21
; Sequence 21, Application US/11018868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guilio
; APPLICANT: Ratti, Guilio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (FP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
```

```
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-21

Query Match          97.4%; Score 5131.5; DB 11; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 MQTSFHFFFLSMILAYSCCSLNGGGYAAAEIWPQGIYDGETLTVSPFYTYIGDPSGTTVF 60
Db 1 MQTSFHFFFLSMILAYSCCSLNGGGYAAAEIWPQGIYDGETLTVSPFYTYIGDPSGTTVF 60
QY 61 SAGELTLKLNLDNSIAALPLSCFQNLGSPFVLRGHSITFENIRTSNNGAALSAAADGL 120
Db 61 SAGELTLKLNLDNSIAALPLSCFQNLGSPFVLRGHSITFENIRTSNNGAALSAAADGL 120
QY 121 FTIEGPKELSPSCNSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180
Db 121 FTIEGPKELSPSCNSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180
QY 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFQENTAQDGGACQVTSFSAMANEAPAFIAN 240
Db 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFQENTAQDGGACQVTSFSAMANEAPAFIAN 240
QY 241 VAGVGGGIAAIVQDQGGQVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
Db 241 VAGVGGGIAAIVQDQGGQVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
QY 301 NGKTLFLNNVSPVYIAAKQPTSGQASNTNNYDGGGAIIFCKNGAQ-AGNNSGVSFPD 359
Db 301 NGKTLFLNNVSPVYIAAEQPTNGQASNTSDNYDGGGAIIFCKNGAQAGNNSGVSFPD 360
QY 360 GEGVVPSSVAAGKGGAIIYAKKLSVANCQVPQFLNNTANDGGAIIYLGESGELSADYG 419
Db 360 GEGVVPSSVAAGKGGAIIYAKKLSVANCQVPQFLNNTANDGGAIIYLGESGELSADYG 420
QY 420 DIIFDGNLKRKTAKENAADVNGTVSSQAIISMGSGGKITTLRAKAGHQILFNDPIEMANGN 479
Db 420 DIIFDGNLKRKTAKENAADVNGTVSSQAIISMGSGGKITTLRAKAGHQILFNDPIEMANGN 480
QY 480 NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSOTGG 539
Db 480 NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSOTGG 540
QY 540 SLYNEAGSTWDFVTPQPPQPPAAANQLITLSNHLSSLANNNAVTPNPPAQDSHP 599
Db 540 SLYNEAGSTWDFVTPQPPQPPAAANQLITLSNHLSSLANNNAVTPNPPAQDSHP 600
QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 659
Db 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 660
QY 660 GNEMPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVVASIVPNSLWGSILDI 719
Db 660 GNEMPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVVASIVPNSLWGSILDI 720
QY 720 RSAHSAIOASVDGRSVCYCRGLWVSGVSNFFYHDDRDLGGQYRIISGGYSLGANSYFGSSMF 779
Db 720 RSAHSAIOASVDGRSVCYCRGLWVSGVSNFFYHDDRDLGGQYRIISGGYSLGANSYFGSSMF 780
QY 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVGFQGNQHMT 839
Db 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVGFQGNQHMT 840
QY 840 SYTFABESDVRWDDNCLVGEIGVGLPIVITPSSKLYLNELRPFVQAEPFSAHDESFTREGD 899
Db 840 SYTFABESDVRWDDNCLVGEIGVGLPIVITPSSKLYLNELRPFVQAEPFSAHDESFTREGD 900
QY 900 QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHOBTW 959
Db 900 QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHOBTW 960
```

```
QY 960 TTDAPLHARHGVVVRGSMYASLTNIEVYGHGRYEVDRASRGVGLSAGSRVRF 1012
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 961 TTDAPLHARHGVVVRGSMYASLTNIEVYGHGRYEVDRASRGVGLSAGSRVRF 1013
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
US-11-103-957-7
; Sequence 7, Application US/11103957
; Publication NO. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Veriant, Vincent Georges Christian Louis
; FILE OF INVENTION: Vaccine Composition
; TITLE OF INVENTION: B45261
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-7

Query Match 14.6%; Score 771.5; DB 11; Length 878;
Best Local Similarity 26.5%; Pred. No. 2.5e-47;
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

QY 53 DPGSTTVFSGAGELTKNLNLSIAALPLSCFNL-----GSFTVLGRGSLTPE 101
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 DPLGETA-----LLTKNPNHVCTFPEDCTMESLFPALCAHASQDDPLVYLGNSYCFVWS 80
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 102 NIRTSTNGAALNSAADGLFTIEGFKELSFNCNSLLAVLPAATTNKGSTPTTTTSPN 161
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 KLHITDPKEALFKEGD--LSIQNFRFLSFTDCSS-----KESSPS- 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 162 GTIYSKT-DLLLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVFOENTAAQDGCAC 220
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 -IIHQKNGQLSLRNGSMSPCRNHAEGSGAIGADAFSLQHNLYLFTAFENSCKNGGAI 178
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 221 QVYTSFSAMANEAPIAFVANVAGVGGGIAAQQDGGQGVSSSTSTEDPVVPSRNTAVEP 280
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 QAQT-FSLSRNVSPISFARNRADLNGAICC----- 208
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 281 DGNVARVGGGIYSGVNVAPLNGKTLPLNNVAGPVYIAAKQPTSGQASNTSNYGGGAI 340
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 209 -----SNLICSGNV-----NLPF-----TGNSATNGAI 233
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 341 FCXKGAQAGNNSGVSFDEGVPVFFSNVAAGKGAIIYAKLISVANCVPQVPLRNAND 400
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 234 CCI--SDLNTSEKGLSLACNQETLFPASNGAKGKGAIIYAKHMYLRYNGVPVFINNSAKI 291
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 401 GGAIYLGESGELSADYDIIIPDGNLKRRTAKENAADVNGVTVSSQAISMGSGKITTLLR 460
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 292 GGAIAIQSGSLSLAGEGSLFPQNSQRTSDQGLVR-NAIYLEKDAI-----JSSLE 343
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 461 AKAGHQLLFNDPI-EMANGNNPAQSSKLLKINDGEGYTG-----DIVFA----- 504
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 344 ARNG-DILFDPPIVQESSESKEPPLSSLSQVSTPTATASPLVIQTSANRVSIFSSERL 402
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 505 -----NGSSTLYQNVITIEQGRIVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 403 SEBEKTPDNLTSQLOQPIELKSGRLVKDRAVLSAPSLSDPQALLIMEAGTS----- 455
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 556 PPOQPPAANQLITLNLHLSSLLANNVTPNPPAQDQSHPAVIGSTTAGSVTISGP 615
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 456 -----LKTSSDLKATLSIPLHSL-----DTEKSVTIHAP 485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 616 -----IPFEDL-DDTAYDRYDMLGSKNQKINVLKQLGCTKPPANAPSDLTL--GNEMPKY 666
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 486 NLSIQKIFLSNGSDENFYENVELLSSKQON-NIPLTL-----SKEQSHLHLPDGNLSHP 539
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 667 GYQGSWKLAWDPNTANNPYTLKATWTKTGYNPGPVERVASLVPNSLWGSIIIDIRSAHSAI 726
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 540 GYQGDWTFESW--KQSDGEG-HSLIANWTPKNYVPHPEROSTLVANTLWNTYSDMQAVQSMI 596
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 727 QASVDGRSYCRGLWVSGVSNFFY-HDRDALG--QGVRYISGGVSLGANSY-FGSSMFGLAF 783
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 597 NTIAHGAYLFGTWSGSAVSNLFYAHDSGKPIDNWHHRSLGYLFGISTHSLDDHDFCLAA 656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 784 TEVFGSRKDYVVCRRSNHHACIGSVLSTQOALCGSYLFGDAPIRASYFGNGHMTSY-T 842
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 657 GQLLGKSDSFITSTE-----TTSIATVQAOLATPLM-KISAQACYNESIHETKTKYS 710
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 843 PAEESDVRWNNCLAGEICAGLPIVITPSKLYNELRPFVQAEFSYADHESFTEGDOAR 902
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 711 FSKEGFGSWHSAVSVSGEVCASIPVNSGSLF-SSFSIFSKLQGFSGTQDGFEESSGEIR 769
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 903 AFKSGHLLNLVSPGVKFDRCSTHNPKNYSFMAAYICDAYRTISGTETTLTLLSHQETWTD 962
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 770 SFSASSFRNISLPMGITTEKKSQKTRNYYFFILGAYIQDLKRDVESGPVVLLKNAVSWDAP 829
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 963 AFHLARHGVVVRGSMYASLTNIEVYGHGRYEVDRASRGVGLSAGSRVRF 1012
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 830 MANLDSRAYMFLRITNQRAL-HRLQTLNVSIVYLRGQSHSYSLDLGLTYRPF 878
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
US-11-018-868-23
; Sequence 23, Application US/11018868
; Publication NO. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Guido
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-23

Query Match 14.6%; Score 771.5; DB 11; Length 878;
Best Local Similarity 26.5%; Pred. No. 2.5e-47;
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

QY 53 DPGSTTVFSGAGELTKNLNLSIAALPLSCFNL-----GSFTVLGRGSLTPE 101
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 DPLGETA-----LLTKNPNHVCTFPEDCTMESLFPALCAHASQDDPLVYLGNSYCFVWS 80
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 102 NIRTSTNGAALNSAADGLFTIEGFKELSFNCNSLLAVLPAATTNKGSTPTTTTSPN 161
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 KLHITDPKEALFKEGD--LSIQNFRFLSFTDCSS-----KESSPS- 119
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 162 GTIYSKT-DLLLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVFOENTAAQDGCAC 220
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 -IIHQKNGQLSLRNGSMSPCRNHAEGSGAIGADAFSLQHNLYLFTAFENSCKNGGAI 178
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 221 QVYTSFSAMANEAPIAFVANVAGVGGGIAAQQDGGQGVSSSTSTEDPVVPSRNTAVEP 280
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 QAQT-FSLSRNVSPISFARNRADLNGAICC----- 208
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 281 DGNVARVGGGIYSGVNVAPLNGKTLPLNNVAGPVYIAAKQPTSGQASNTSNYGGGAI 340
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 209 -----SNLICSGNV-----NLPF-----TGNSATNGAI 233
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 341 FCXKGAQAGNNSGVSFDEGVPVFFSNVAAGKGAIIYAKLISVANCVPQVPLRNAND 400
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 234 CCI--SDLNTSEKGLSLACNQETLFPASNGAKGKGAIIYAKHMYLRYNGVPVFINNSAKI 291
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 401 GGAIYLGESGELSADYDIIIPDGNLKRRTAKENAADVNGVTVSSQAISMGSGKITTLLR 460
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 292 GGAIAIQSGSLSLAGEGSLFPQNSQRTSDQGLVR-NAIYLEKDAI-----JSSLE 343
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 461 AKAGHQLLFNDPI-EMANGNNPAQSSKLLKINDGEGYTG-----DIVFA----- 504
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 344 ARNG-DILFDPPIVQESSESKEPPLSSLSQVSTPTATASPLVIQTSANRVSIFSSERL 402
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 505 -----NGSSTLYQNVITIEQGRIVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 403 SEBEKTPDNLTSQLOQPIELKSGRLVKDRAVLSAPSLSDPQALLIMEAGTS----- 455
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 556 PPOQPPAANQLITLNLHLSSLLANNVTPNPPAQDQSHPAVIGSTTAGSVTISGP 615
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 281 DGNVAVGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKOPTSQASNTSNYYDCGAI 340
Db 209 -----SNLCSGNV-----NPLFF-----TGNSATNGGAI 233
QY 341 FCKNGAAGSNNSGVSFGDGBGVFFSNVAGKGGAIYAKKLSVANGCPVQFRLNIAND 400
Db 234 CCI--SDLNTSEKLSLACHQETLFPASNAKEKGAIYAKHVLRYNGVPSPINNSAKI 291
QY 401 GGAIYLGSGBELSLADYDIIIFDGNLKRKTAKENAAADVNTVSSQAISMGSGKITTLLR 460
Db 292 GGAIAIOGGGSLILAGESVLFPONNSORTSDQGLVR-NAIYLEKDAI-----LSSLE 343
QY 461 AKAGHQIILPNDPI-EMANGNQPAGOSKLLKINDGEGVTG-----DIUVA----- 504
Db 344 ARNG-DILFFDPIVQESSKESPLSPSQAASVTSPTPATAPLVIQTSANRSVIFSSRL 402
QY 505 -----NGBSTLYQNTVIEGRVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
Db 403 SEBEKTPNLSLOQLOPIELKSGRLVLDKRAVLSAPLSQDPQALLIMEAGTS----- 455
QY 556 PPQOPPAANQLITLSNLHLSLSSLANNAVNTPTNPPAQDHPAVIGSTTAGSVTTISGP 615
Db 456 -----LKTSBOLKATLSILPHSL-----DTEKSVTIHAP 485
QY 616 -----IPEEDL-DDTAYDRYDMLGSNQKINVLKQLGTCKPANAPSDLTL--GNEMPKY 666
Db 486 NLSIQKIFLSNGSDENFYENVLLSKQN-NIPLLT-----SKBQSHLPLDGNLSHF 539
QY 667 GYQGSWKLAWDNTANNPFTLAKATWTKTGYNPGPERSVASLPNSLWGSILDIRSAHAI 726
Db 540 GYQGDWTSW--KQSDGEG-HSLIANWTPKNTVYHPEROSTLVANTLWNTYSDMQAVQSMI 596
QY 727 QASVDGRSRYCRLMWSGVSNFFY-HDRDALG-QGYRIYSGGYSLGANSY-FGSSMFGIAP 783
Db 597 NTIAHGGAYLPQWGSASVNLFYAHDSSGKPTDNHHRSLGYLFGISTHSLDHSFCLAA 656
QY 784 TEVFGSRKDYVVCNRNHHACIGSVLSTQOALCGSYLFGDAFIRASYFGNQHKITSY-T 842
Db 657 QOLLKSSDSFITSFE-----TTSYIATVOQLATPLM-KISAQACYNESHEILKTKYRS 710
QY 843 PAESDVEDWNNCLAGIAGLPIVITPSKLYNLRLPFPVQAEFSYADHESFTEBGDQAR 902
Db 711 FSKEGFGSHVAVSGEVACSIPIVNSGGLF-SSFSIFSKLQGSFGTQDGFESSGEIR 769
QY 903 AFKSGHLLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHOBTWTD 962
Db 770 SPASASSPRNLSLPMGSIITPEKKSQKTRNYFFLGYIQLKRDVBSGPPVLLKNAVSWDAP 829
QY 963 AFHLARHGVVVRGSMYASLTSNIEVYHGRIYDRDASRGYLSAGSRVRP 1012
Db 830 MANLDSRAYMFLTMORAL-HRLQTLNVSYYLRGQSHSYSLDLGTTTYRF 878
```

## RESULT 5

```
US-11-103-957-13
; Sequence 13, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; TITLE OF INVENTION: Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
```

```
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-13
```

```
Query Match 12.3%, Score 649.5; DB 11; Length 964;
Best Local Similarity 25.8%; Pred. No. 1.8e-38;
Matches 260; Conservative 155; Mismatches 389; Indels 205; Gaps 43;
```

```
QY 84 NLLGSPTVLGRGSLT---PENIR-----TSTNGAALSNSAADGLFTIEGKELSPS 132
Db 39 SLSNKISLTGDTNHLTCYLDLNLRYLAILQKTPEGAAV--TITDYSLSPDFTQKEGIYF 96
QY 133 NCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLLNNE--KFSFVSNLVSGD-- 188
Db 97 AKN-----LTPESGGAIGYASPNSTVEIRDTI---GPVFNENTCCRLFTWRNPYADKI 149
QY 189 --GGAIDAKSLTVQGISKLCVFOBNTAQADGGACQVVTFSAMANEAPIAPVANV----- 241
Db 150 REGGAIHAQNLINHNHDVVGMKNFSGVQGAISTANTFVVSQSCFLFMDNICIQTN 209
QY 242 -AGVRGGIAAVQDQGVSSSTSTEDPVVSPSNTAVEFDGNVARVGGGIYS----- 293
Db 210 TAG-KGGAIIYA-----GTSNSFE-----SNCDLFFINNACCAGGAIIPSPICSLTG 254
QY 294 -YGNVAFNLNGKTLFLNNVASPVYIAAKOPTSQASNTSNYYGCGAIFCKNGAAGSNN 352
Db 255 NRGNIIVFNN--RCFKN-----VETASSEASDGGAIKVTRLDVTGN- 294
QY 353 SGSVDFGEGVVPFSSNVAAGKGAIIYAKKLSVANGCPVQFRLNIAND-GGAIYLGSGGE 411
Db 295 -----RGRIPPSDNIKNYOGAIYAPVTVLVNDGPTVFYFINNIANNKGAIYIDGTSN 346
QY 412 LSLSADYDIIIFDGNLKRKTAKENAAADVNTVSS-----QAISMGSGKITTLLRAKHQ 466
Db 347 SKISADRHAIIFENI-----VTNVTNANGTSTGANPPRRAITVASSSGEILLGAGSQN 402
QY 467 ILFNDPIEMANGNNQPAQSKLLKINDGEGVTGDIVPAN--GSSTLYQN-----V 514
Db 403 LIFFDPIEVSN-----AGSVSFNKEADQTSVSVFSGATVNSADFHQNLQTKTAPL 455
QY 515 TIBOGRIVLREKAKLSVNSLSQTGSLYMBAGS-----TWDFVTPOPPQPPAANQL 566
Db 456 TLSNGFLCIEDHAQLTVNRFTQTGGVVVSLGNAGVLSCKYKNGTGD-----SAGNAS 505
QY 567 ITLSNLHLSLSSLLANNA-----VTNPTNPPAQDSDHPAVIGST-TAGSVTISGPIPEDL 621
Db 506 ITLKHGLNLSSILKSGAEIPLLMVEPTN--NSNNYTADTAATFSLSDVKLS---LIDDY 560
QY 622 DDTAYDRYDMLGSNQKINVLKQLGTKP--PANAPSDLTLGNE-----MPKYGYQGS 671
Db 561 GNSPYESTD-----LTHALSSQPMLSISEASDNQLOSENIDFSGLVNPHYGQQL 610
QY 672 WKLAW-----DPTNANNPYTLKATWTKTGYNPGPERSVASLPNSLWGS- 715
Db 611 WTMGWAKTQDPEPASSATITDPQKANRFHRTLTLTLPAGYVPSPKHRSPLIANTLWGNM 670
QY 716 ILDIRSAHSAIOASVDGRSVCRLWVSGVSNFFVHRRDALCGQYRIYSGGYSLG---ANS 772
Db 671 LLATESLKNSAELTSPSGHPFW-GITGGGLGMVYQDPREHNPGHMSSSGYSAGIAQQT 729
QY 773 YFGSSMFGLAFTVEFGR-8KDYVVCNRNHHACIGSVYLSLSTQQA-----LCSYSLFGDAFI 826
Db 730 HTFSLKFSQTYTKLNERYAKNNV--SSKNYSQCGEMLPQLQGBGLLTKLVGLIYSGDHNC 787
QY 827 RASYFGNQHKITSYTPABESDVWMDNNCLAGEIAGLPIVITPSKLYNLRLPFPVQAEF 886
Db 788 HHFYTQG-ENLTSQGTFRSQT-----MGAVFFDLPMKPFGSTHILT--APFLGALG 836
QY 887 SYADHESPTBEGDQARAFKS-GHLLNLNLSVPVGVKFDRCSSSTH-PNKYSFMAAYICDAYRT 944
```

Db 837 IYSSLSHTEVGGAYPRSTKTPPLINLVPIGVKGSFNNATHRPQAWTVLAYQPVLYRQ 896  
Qy 945 ISGTETLLSHQETWTTDAFHARHGCVVVRGSMYAS-----LTSNIEVYG 989  
Db 897 EPGIAAQLASKGIWFGSGSPSSRHAMSYKISQQTQPLSLWTLHFQYHG 945

RESULT 6  
US-11-018-868-19  
; Sequence 19, Application US/11018868  
; Publication No. US20060034871A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Giulio  
; APPLICANT: Bonci, Alessandro  
; APPLICANT: Pincio, Oretta  
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis  
; FILE REFERENCE: 002441.00099 (P23152.001)  
; CURRENT APPLICATION NUMBER: US/11/018,868  
; CURRENT FILING DATE: 2004-12-22  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 19  
; LENGTH: 964  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-11-018-868-19

Query Match 12.3%; Score 649.5; DB 11; Length 964;  
Best Local Similarity 25.8%; Pred. No. 1.8e-38; Indels 205; Gaps 43;  
Matches 260; Conservative 155; Mismatches 389;

Qy 84 NLGSPVLGRGHSLT---FENIR-----TSTNGAALSNSAADGLFTIEGFKELSPS 132  
Db 39 SLNKSISLTGDTNLTNLCYDLNRYLAILQKTPNEGA--TITDLSFFDTQKEGIYP 96  
Qy 133 NCNLSLAVLPAATNKGSTPTTTPSNGTIYSKTDLLLNNE--KFSFYSLNLSGD-- 188  
Db 97 AKN-----LTPESGGAIGASPNSTPVEIRDITI---GPVIFENNTCCRLFTWRNPFYADKI 149  
Qy 189 ---GGATDAKSLTVQGISKLCVFQENTQAADGGACQVVTFSAMANEAPAFVANV----- 241  
Db 150 REGGALHAQNLINHNHVVGVFMKFSYVQGGAI StantonFVVSNSQSCFLFMDNICIQTN 209  
Qy 242 -AGVRGGAAVQDGGQGVSSSTSTEDPVVFSFRNTAVEFDGNVARVGGIYS----- 293  
Db 210 TAG-KGGAIA-----GTNSFE-----SNCDLFFINNACCAGGAIFPICSLTG 254  
Qy 294 -YGNVAFNLNGKTLFLNNVASPVYIAAKOPTSQASNTSNYGDGGAIFCKNGAQAQGSNN 352  
Db 255 NRCNIVPYNN--RCFKN-----VETASSEASDGGAIKVTTRLDVTGN- 294  
Qy 353 SGVSFDEGEVGFSSNVAAGKGAIAKLVANCGVPQFPLNIAND--CGAIVLGESGE 411  
Db 295 -----RCRIFSDNITKNYGGAIYAPVITLVDNGPTYPFINNIANNKGGAIYIDGTSN 346  
Qy 412 LSLSDYGDIIIPGNLKRKAKENAADVNGVTSS-----QAISMGSGKITTILRAKAGHQ 466  
Db 347 SKISADRHAIIFENI-----VTNVTNANGTSTANPPRRNAITVASSSGEILLGASSQN 402  
Qy 467 ILFNDPTIEMANGNNQPAQSSKLKINDGEGYTDIVFAN---GSSTLYQN-----V 514  
Db 403 LIYDPIEVSN-----AGVSVSFNKEADQTSVVFSGATVNSADPHQNLQTKTPAPL 455  
Qy 515 TIEGRVILREKAKLVNSLSQTCGSLYMEAGS-----TWDFTVPPQPPQPPANQL 566  
Db 456 TLSNGFLCIEDHAQLTNRFTQTGGVSVSLGNAGVLSYKNGTGD-----SASNAS 505  
Qy 567 ITLSNLHLSLSLILANNA-----VTNPPTNPPAQDSHPAVIGST--TAGSVTISGPIPEDL 621  
Db 506 ITLKHGLNLSLILKSGABEPLLVWEPTN--NSNNYADTAATFSLSDVKLS-----LIDDY 560

Qy 622 DDTAYDRYDMLGNSQKINVLKQLQGTKP--PANAPSDLTJLGN-----MPKGYGQS 671  
Db 561 GNSPYESTD-----LTHALSSQPMLSISEASDNQLOSENIDFSGNLVNPYHGMQGL 610  
Qy 672 WKLAW-----DPNTANNPPTLKATWTKVGNPGRVASLVPNLSLWGS- 715  
Db 611 WTWGAKTQDPEPASSATITDPQANRPHRTLLLTWLPAGVYVSPKHSPLIANTLWGM 670  
Qy 716 ILDIRSAHSAJQASVDGRSYCRGLWVSGVSNFFYHDDRDLAQGYRIISGGYSLG---ANS 772  
Db 671 LLATESLKNASIELTSPGHPFW-GITGGGLGMVYQDPRENHPGFTHRSSGYSAGMAGQT 729  
Qy 773 YFGSSMFLAFTVEYFGR-SKDYVVCVRNHHACISGVYLSLSTQQA-----LCSYLPEDAFI 826  
Db 730 HTFSLKFSQTYTKLNERVAKNV--SSKNYSQCGEMLFSLQEGFLLTKLVGLVSYGDHNC 787  
Qy 827 RASVFGNGHMKTSYTPABESDVRDNNCLAGEIGAGLPIVITPSKLVNLNLRPPVQAEF 886  
Db 788 HHFTYQG-ENLTSQGTFRSQT-----MGAVFFDLPMKPFPGSTHILT--APFLGALG 836  
Qy 887 SYADHESFTEGDOARAFKS-GHLLNLSVPVGVKFDRCSSTH-PNKYSFMAAYICDAYRT 944  
Db 837 IYSSLSHTEVGGAYPRSTKTPPLINLVPIGVKGSFNNATHRPQAWTVLAYQPVLYRQ 896  
Qy 945 ISGTETLLSHQETWTTDAFHARHGCVVVRGSMYAS-----LTSNIEVYG 989  
Db 897 EPGIAAQLASKGIWFGSGSPSSRHAMSYKISQQTQPLSLWTLHFQYHG 945

## RESULT 7

US-11-103-957-41  
; Sequence 41, Application US/11103957  
; Publication No. US20050281847A1  
; GENERAL INFORMATION:  
; APPLICANT: Berthet, Francois-Xavier Jacques  
; APPLICANT: Lobet, Yves  
; APPLICANT: Poolman, Jan  
; APPLICANT: Verlant, Vincent Georges Christian Louis  
; TITLE OF INVENTION: Vaccine Composition  
; FILE REFERENCE: B45261  
; CURRENT APPLICATION NUMBER: US/11/103,957  
; PRIOR FILING DATE: 2005-04-12  
; PRIOR APPLICATION NUMBER: US/10/467,534  
; PRIOR FILING DATE: 2004-02-03  
; PRIOR APPLICATION NUMBER: PCT/EP02/01356  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: GB 0103169.9  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 1016  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-11-103-957-41

Query Match 11.7%; Score 615; DB 11; Length 1016;  
Best Local Similarity 24.2%; Pred. No. 6.1e-36;  
Matches 280; Conservative 163; Mismatches 402; Indels 312; Gaps 56;

Qy 3 TSPHKFELSLAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFPVTIVIGDPSGTVFSA 62  
Db 8 TSP--CFLACLCSYSY-----GFSS--PQVL-----TPNVTTFP-----KGGDVLN 46  
Qy 63 GELTLKXNL-----DNSIAALPLSCFQNLGSLFTVLGRHSLTFTENIRTSNGLAALSNA- 116  
Db 47 GDCAFVNVYAGANGSI-----ISANGD---NLITQONHTLSP-----TDSQGVLPQNYAF 95  
Qy 117 --ADGLFTTEGFKELSPS---NCN-----SLLAVLPAA 144  
Db 96 ISAGETLTLKDFSLFSLMFSKNVSCGRKMISGKTVSISGAGEVIFWDSNVGYSPLSIVPAS 155

```
QY 145 TTNGSQTP-----TTSTPSTYKTDLLLLLNKEKSFYSNLVSGDGAIDAK 195
Db 156 TPTPPAPAPAAASSLSPTVSDARKGSIFS-----ET 189
QY 196 SLTVQIGISKLCVFOBNTAQADGACQVVTSPSAMANEAPIAFVANVAGVGGIAAVQDG 255
Db 190 SLEISGVKGVMP-----DNNAGNFGTVPRGNSN-----NNAAGSGSG-----227
QY 256 QQGVSSSTSTEDPVVFSRNTAVEPDGNVARGGIIYGVNVAFLNN-GKTLFLNVA-S 313
Db 228 -----SATTPSFVKCKG-KVSTDNVASCQGVVYKGTVLPKDNBEGGIFPRGNTAYD 280
QY 314 PVYIAAQPTSGQASNTNNYGDGAIKCK-NGAQAGSNNSGVSFD-----GEG-----362
Db 281 DLGILA---ATSRDQNTET---GGGGVICSPDDSVKFEKNGKSVFDYFNPAKGRGGSILT 335
QY 363 -----VVFSSNVAAGKGAIYAKKLSVA-NCGPVQFLRNANDGGAIIYL-----G 407
Db 336 KEPSLVADDSVVFSSNNTAEKGGAIYAPTIDISTNGGSIILFERNRAEGGAIKVSSESSG 395
QY 408 ESELSLSADYDGIIFDGNL--KRTAKENAADV--NGVTSSQAISSMSGGKITTILRAKA 463
Db 396 STGNLTLSASDGDIVFSGNMTSDRPGERSAARILSDGTTVS-----LNASG 441
QY 464 GHQILFNDPIEMAN---GNNQPAQSSKLL--KINDGEGYTDIVFA-----N 505
Db 442 LSKLIFDYPVQNNASAGASTPSPSSMPGAVTINOSGNSVIFTAESLTPSEKLOVLN 501
QY 506 GSSTLYQNVTIEQRIVLREKAKLSVNSLSQTSGLSMEAGSTWDFVTPQPPQPPAANQ 565
Db 502 STSNFPGALTVSGGELVVTEGATLTGTTITATSRVTLGSGASLSAVA-----GAANN 554
QY 566 --LITLSNLHLSLSSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTIGSIPFEDLDD 623
Db 555 NYTCTVSKLGDLESFLTPNYKT-----AILGAD--GTVTVNSGSTLDLVM 599
QY 624 TAYDRYD---WLGSNQKINVLKLGKTPPANAPSDLTGN-----EMPKYGVQGSWKL 674
Db 600 SEAEVYDNLFPVGS-----LTIPFVTLSSSSASNGVTNKSVTINDADAAHYQGSWSA 653
QY 675 AW-----DPTANNGPVTLLKATWTKTG---YNPGERVASLVPNSLWGSIL 717
Db 654 DWTXPPLAPDAKGMVPTNNK---TLYLTVRPASNYGEYLDPOKGEVFNLSWAGS 709
QY 718 DIRSAHSAIQASVDGR--SYCRGLMWVG--VSNFYHDDRDLGQGYRIISGGY-SLGANS 772
Db 710 ALRFTPTNGLKEHYSRDGVFVASLHALGDYILNVTQDDRD---GFLARYGGFOATAASH 765
QY 773 YFGSMFLATEVPGRSKDYVVCESNKH-----ACIGSVYL---STQOALCGSYLPGDA 824
Db 766 YENGSIQVAFQGLYGQTKSRMYYSKODGNMTLSCFGRSYVDIKGTETVN-----816
QY 825 FIRASYPGNQHMKTSY-----TPAESEDRVWDNNCLAGETGAGLPIV--ITPSKLYLN- 876
Db 817 YWETAYGVSVRMTQYFNDKQKEDHSKCHWNHNYAFVGAENFLEYCIPRQFARD 876
QY 877 -ELRPPVQAFSADHESFTEEGDQARAFKSHLLNLSPVGVKFDRCSTH---PNKYS 932
Db 877 YELTGPMRPEMAGCWSSTRBTGSLTRYFARGSGHNSLPIGIVAHAVSHVRSPPSKLT 936
QY 933 PMAAYICDAYTISGTETTLTSHOBTWTDAPHLARHGVVVRGMYASLTSNIEVYHGR 992
Db 937 LNMGYRPDIWRTVPHCNMBIITANGVKTIQGSPLARH-----APFLEVHDTLTHHFR 990
QY 993 YEYRDASRGYGLSAGSR 1009
Db 991 -----AYMNYSLDARR 1002
```

RESULT 8

US-11-018-868-22

; Sequence 22, Application US/11018868

; Publication No. US20060034871A1

```
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 1016
; TYPE: PRN
; ORGANISM: Chlamydia trachomatis
US-11-018-868-22
```

Query Match 11.7%; Score 615; DB 11; Length 1016;  
Best Local Similarity 24.2%; Pred. No. 6.1e-36;  
Matches 280; Conservative 163; Mismatches 402; Indels 312; Gaps 56;

```
QY 3 TSPHKFPLSMILAVSCCSLNGGGVAAEIMVPOGIYDGETLTVSPYTVIGDPSGTTVPSA 62
Db 8 TSP--CFCLACLSYSY-----GPASS---PQVL---TPNVITPF-----KGDDVVLN 46
QY 63 GELTKLNL-----DNSIAALPLSCFNLGASFTVLGRGHSITFENIRTSNTGAALSNSA- 116
Db 47 GDCAFNVYAGAENGSI-----ISANGD---NLITIGQNHLSF---TDSQGPVLQNVAP 95
QY 117 --ADGLFTIGFKELSPS---NCN-----SLAVLPAA 144
Db 96 ISAGETLTLKDFSLMPSKNVSCGEKMIKSTVISISGAGEVIFMDNSVGSYSPLSIVPAS 155
QY 145 TTNGSQTP-----TTSTPSTYKTDLLLLLNKEKSFYSNLVSGDGAIDAK 195
Db 156 TPTPPAPAPAAASSLSPTVSDARKGSIFS-----ET 189
QY 196 SLTVQIGISKLCVFOBNTAQADGACQVVTSPSAMANEAPIAFVANVAGVGGIAAVQDG 255
Db 190 SLEISGVKGVMP-----DNNAGNFGTVPRGNSN-----NNAAGSGSG-----227
QY 256 QQGVSSSTSTEDPVVFSRNTAVEPDGNVARGGIIYGVNVAFLNN-GKTLFLNVA-S 313
Db 228 -----SATTPSFVKCKG-KVSTDNVASCQGVVYKGTVLPKDNBEGGIFPRGNTAYD 280
QY 314 PVYIAAQPTSGQASNTNNYGDGAIKCK-NGAQAGSNNSGVSFD-----GEG-----362
Db 281 DLGILA---ATSRDQNTET---GGGGVICSPDDSVKFEKNGKSVFDYFNPAKGRGGSILT 335
QY 363 -----VVFSSNVAAGKGAIYAKKLSVA-NCGPVQFLRNANDGGAIIYL-----G 407
Db 336 KEPSLVADDSVVFSSNNTAEKGGAIYAPTIDISTNGGSIILFERNRAEGGAIKVSSESSG 395
QY 408 ESELSLSADYDGIIFDGNL--KRTAKENAADV--NGVTSSQAISSMSGGKITTILRAKA 463
Db 396 STGNLTLSASDGDIVFSGNMTSDRPGERSAARILSDGTTVS-----LNASG 441
QY 464 GHQILFNDPIEMAN---GNNQPAQSSKLL--KINDGEGYTDIVFA-----N 505
Db 442 LSKLIFDYPVQNNASAGASTPSPSSMPGAVTINOSGNSVIFTAESLTPSEKLOVLN 501
QY 506 GSSTLYQNVTIEQRIVLREKAKLSVNSLSQTSGLSMEAGSTWDFVTPQPPQPPAANQ 565
Db 502 STSNFPGALTVSGGELVVTEGATLTGTTITATSRVTLGSGASLSAVA-----GAANN 554
QY 566 --LITLSNLHLSLSSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTIGSIPFEDLDD 623
Db 555 NYTCTVSKLGDLESFLTPNYKT-----AILGAD--GTVTVNSGSTLDLVM 599
QY 624 TAYDRYD---WLGSNQKINVLKLGKTPPANAPSDLTGN-----EMPKYGVQGSWKL 674
Db 600 SEAEVYDNLFPVGS-----LTIPFVTLSSSSASNGVTNKSVTINDADAAHYQGSWSA 653
```





```
RESULT 10
US-11-018-868-18
; Sequence 18, Application US/11018868
; Publication No. US2006003487A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-18

Query Match 11.3%; Score 597.5; DB 11; Length 1531;
Best Local Similarity 24.1%; Pred. No. 2e-34;
Matches 293; Conservative 153; Mismatches 474; Indels 295; Gaps 54;

Qy 17 SCCSLNGGGYAAABIMV-----PQGI-YDGETITVSPVTVIGDPSPGTTVFAGBELTLKNL 70
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 393 STACLGGAIAAQEIVSIQNNQAGISFEGKASFGG-----GIACGSPSSAGGASVLGTI 447
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 71 D-NSIAALPLS---CFGNLLGSFTVLGRGHSHTPEINRTSTNGAALSNSAADGLFTIEG 125
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 448 DISKMLGAISSRTCTTSDLGQMEYQG--GGALFGENISLSNAGVLT-----494
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 126 FKE-----LSFSNCSNLLAVPAATTKGSGTPTTSTPSNGTIYSKTDLLLLNNEKFSFYS 182
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 495 PKDNIVKTFASNGKILGGGAILATGKVEITNNSSEGISFTGN--ARAPQALPTQEEFPLPS 552
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 183 N-----LVSG--DGAIDAKSLTVGIGSKLCVFPQENTQAQD-----GGACQVVT 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 553 KKEGRPLSSGGGAILGREVAILH--NAAVVFQENRLQCSSEEAATLLGCGCGGAVHGM 611
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 225 SFSAMANAAPATFANVA---QVRGGGIAAQDQGGQSSSTSTEDPVVPSRNTAVSPD 281
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 612 STSIVGNS--VAFGNVYAGQVSGGALLS-----KTVLAGNSVDVFS 654
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 282 GNVARVGGGIY--SYGNVAFLLNGKTL-----PLN 309
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 655 RNIALGGGALQASRGNCGLVDNGVYLPFDNRGRVYGGAIISCLRGDVPVIGSKGRVFPKD 714
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 310 NVASPVYTAAK-----OPTSGQASNTSNY-----GD-----336
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 715 NIATRLVYEETVEKVEEPAPEQKDNNELSFLGRAEQSFITAANQALFASDEDGDLSPES 774
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 337 -----GGALFCK-----NGAQAGSN-----SGSV-----SPDGE 361
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 775 SISSEELAKREACAGAIKAKRVRIVDQEAUVVSNFSDIYGGAIFTGSLREEDKLDGQ 834
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 362 -----GVVFPSSNVAAGK-----GGAIYAKKLSVA--NCGVPVQFLRNIANDGGA 403
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 835 IPEVLISGNAGDVVFGNSKRDHLPHTGGGAICTQNLTIISQNTGNVLFYNNVACSGGA 894
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 404 IYLGESGELSADYGDIIIFQNLKRTAKENAAVNGVTVSSQAISM--GSGGKITTLRAK 462
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 895 VRIEDHGNVLLFAFGDITVFKGNSFRAQ-----GSDAIYPAGKESHITALNAT 943
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 463 AGHQLFLNDPIEMANGNNOQAQSKLLKINDGE--GYTGDIVPANGSSTLYQNVTIEGR 520
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 944 EGHAIIVFDALVFE--NLEERKASVLLINSRENPGYTGSRFLEASKVFPQCIHVQVGS 1001
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 521 IVLREKAKLSVNSLSQ--TGGBSLYBAGSTWDFV--TPQPPQPPAAN-----QLITLSNLHL 574
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1002 LELLNGATLCSYGFQKQDAGAKLVLAAGAKKLIDSGTVPQQGHAIKPEAEISSSBEPG 1061
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

US-11-103-957-47
; Sequence 47, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berchet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 975
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-47

Query Match 10.5%; Score 551; DB 11; Length 975;
Best Local Similarity 24.9%; Pred. No. 2.4e-31;
Matches 279; Conservative 148; Mismatches 405; Indels 290; Gaps 57;

Qy 3 TSPHKFFLSMILAYSCCSLNGGGYAAABIMV-----PQGIYDGETITVSPVTVIGDPSPGT 57
Db 32 TLIPKFLGALIVVAPYSF-----AEWELAISGHKQG--KORDTFTW-----ISSCPBGT 79
```

```
QY 58 TVFSAGELTLK--NLDNSIAALPLSCFGLNLLSGFTVLGRGH--SLTFENIRSTNGAALS 113
Db 80 NYIINRKLILSDPSLKNVSS--GCAFRNLAKISFLGKNSSASIFPKHININGFGAGVF 137
QY 114 NSAADGLFTTIEGFKELSPNCSNLSLAVLPAAATNKGSTPTTSTPSNGTIYSKTDLLLL 173
Db 138 SSSS-----IEFTDLRKLVAF-----GSES-----TGIFTAKEDISPK 171
QY 174 NNEKFSFYNLVSNGDGAIDAKSLTVQGISK-----LCVFOENTAQAAGGACQV 223
Db 172 NNHHIAFRNITKNGGVQ-----LOGDMKGSVFDQORGAIITNN-----QAV 217
QY 224 TSFSAMANEAPAFVANVAGVRGGIAAVODQOQGVSSSTSTEDPVVFSRNTAVEFDGN 283
Db 218 TS--SSMKHSG-----RGAISGDFAGSR-----ILFLNNQOITFEGN 253
QY 284 VARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPT---SGOASNTSN---NYGDG 337
Db 254 SAVHGGAIYN-----KNGLVEFLGN--AGP--LAPKENTTIANGGAIYTSNFKANQOTS 303
QY 338 GAIFCKNGAQAGSNNSGSVSF-----DREGVVPFSSNVAAGKGAIYAKKLVSANCGP 390
Db 304 PILFSQNH-----NKKGGAIYAQYVNLEQNDTIRFEKNTAKEGGAIYTSQCSITAHNT 359
QY 391 VOFLRNIAND--GGATYL--GESGELSLSDYGDIIIPDGN--LKRTAKENAADVNGVTVS 445
Db 360 IIFSDNAAGDLGGAILLEGKPSLTLIAHSGNIAFSGNTMLHTTKASLDHNSILIKE 419
QY 446 QAISMGGSGKITTLRAKAGHQILFNDPIEMANGNNQOASSKLLKINDGEGYT-----G 499
Db 420 APYKI-----QLAANKNHSIHFFDPVMASSSPIQ-----INAPETPTFPSPKG 466
QY 500 DIVP-----ANGSSTLYQNVTIEQRIVLRREKAKLSVNSLSQTGSLYMEAG 546
Db 467 MIVFSGANLLDDAREDVANRTSIFNPQVHLYNGTSLIENGALHIVOSFKQTGRISLSPG 526
QY 547 STWDFVTPQ-----PPOQPPAANOLITLSNLHLSLSLLANNAVTPNTPPPAQDSDHP 599
Db 527 SSLALYTMNSFFHGNISKEPLEINGL--SFGVDISPSNLQAEIRAGNAPLR-----576
QY 600 AVIGSTTAGSVTISGP--IFFEDLDDTA--YDRYDWLGSNQKINVLKQLGTGKPPANAPS 655
Db 577 -----LSGSPSIHDEGLFYENRDTAASPYQMEILLTSDKIVDISK-----TT 620
QY 656 DLTGLNEMPKYQGSWKLAWDNTANNGPY--TLKATWTKTG--YNGPDERVASLVNSLW 713
Db 621 DSLVTNK--QSGPQGAWHFSWQPTNTNTKQILRASWLPTEGYVLESNRVGRVAPVNSLW 678
QY 714 GSTLDIRSA-----HSAIQASVDGRSYCRGLWVSGV-----SNFFYHDDRDLGQ 757
Db 679 STFLLQLTASHNLGDHLNNRSLIPTSY--GVLIIGTGABMSTHSEESFISRLGATGT 737
QY 758 GYRYISGYSILGANSYFGSSMFLAPTEVFGRSKDYVVCRSNHHACIGSVY-----808
Db 738 SIIRLTPSLTSGG---GSHWFEDSP-----VADLPEHITSEGIQVNVGLTHVMG 784
QY 809 -LSTOQALCSYLFGBDAFIRASYFGNQHKMTSYTFABESDVWDNNCLAGEIGAGLPV 867
Db 785 PLTVNSTLCAA--LDHNAVRIC-----SKDHTYV-----KWDTFGMRGTILGASVTF 831
QY 868 ITPSKLYNLRLFPVQAFPSYADHE-----SFTREGDQARAFKSGHLLNLSPVGVKPF 920
Db 832 -----EYDQTMR-----VPSFANIEATNIIQRAFTETGYNPRFSFKTLNIAIPIGIGY 881
QY 921 DRCSSTHPNKSFWMA-----AVICDVRTISGTETTLILSHQETWTD-----APHLARHG 970
Db 882 EFLCUN--SSFALLKGSIGYSRDIKENPSTLAHLAMNDFAWTNGCSVPTSHTLANQ 939
QY 971 VVVRGSMYASLTNSIEYHGRVEYRDASRGYGLSAGSRVRP 1012
Db 940 LILR---YKACSLVITATYINR--EGKNLSN--SLSCGGYGVF 975
```

```
RESULT 12
US-11-018-868-15
; Sequence 15, Application US/11018868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PF23152.001)
; CURRENT APPLICATION NUMBER: US/11/018.868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 975
; TYPE: PRN
; ORGANISM: Chlamydia trachomatis
US-11-018-868-15

Query Match 10.5%; Score 551; DB 11; Length 975;
Best Local Similarity 24.9%; Pred. No. 2.4e-31;
Matches 279; Conservative 148; Mismatches 405; Indels 290; Gaps 57;

QY 3 TSFHKFFLSMILAYSCSLNGGGYAAEIMV-----PQGIYDGETILTVSPYTVIGDPSGT 57
Db 32 TLIPFLGALIVAPYSF-----AEMELAISGHKQG-KORDITFM-----ISSCPEGT 79
QY 58 TVFSAGELTLK--NLDNSIAALPLSCFGLNLLSGFTVLGRGH--SLTFENIRSTNGAALS 113
Db 80 NYIINRKLILSDPSLKNVSS--GCAFRNLAKISFLGKNSSASIFPKHININGFGAGVF 137
QY 114 NSAADGLFTTIEGFKELSPNCSNLSLAVLPAAATNKGSTPTTSTPSNGTIYSKTDLLLL 173
Db 138 SSSS-----IEFTDLRKLVAF-----GSES-----TGIFTAKEDISPK 171
QY 174 NNEKFSFYNLVSNGDGAIDAKSLTVQGISK-----LCVFOENTAQAAGGACQV 223
Db 172 NNHHIAFRNITKNGGVQ-----LOGDMKGSVFDQORGAIITNN-----QAV 217
QY 224 TSFSAMANEAPAFVANVAGVRGGIAAVODQOQGVSSSTSTEDPVVFSRNTAVEFDGN 283
Db 218 TS--SSMKHSG-----RGAISGDFAGSR-----ILFLNNQOITFEGN 253
QY 284 VARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPT---SGOASNTSN---NYGDG 337
Db 254 SAVHGGAIYN-----KNGLVEFLGN--AGP--LAPKENTTIANGGAIYTSNFKANQOTS 303
QY 338 GAIFCKNGAQAGSNNSGSVSF-----DREGVVPFSSNVAAGKGAIYAKKLVSANCGP 390
Db 304 PILFSQNH-----NKKGGAIYAQYVNLEQNDTIRFEKNTAKEGGAIYTSQCSITAHNT 359
QY 391 VOFLRNIAND--GGATYL--GESGELSLSDYGDIIIPDGN--LKRTAKENAADVNGVTVS 445
Db 360 IIFSDNAAGDLGGAILLEGKPSLTLIAHSGNIAFSGNTMLHTTKASLDHNSILIKE 419
QY 446 QAISMGGSGKITTLRAKAGHQILFNDPIEMANGNNQOASSKLLKINDGEGYT-----G 499
Db 420 APYKI-----QLAANKNHSIHFFDPVMASSSPIQ-----INAPETPTFPSPKG 466
QY 500 DIVP-----ANGSSTLYQNVTIEQRIVLRREKAKLSVNSLSQTGSLYMEAG 546
Db 467 MIVFSGANLLDDAREDVANRTSIFNPQVHLYNGTSLIENGALHIVOSFKQTGRISLSPG 526
QY 547 STWDFVTPQ-----PPOQPPAANOLITLSNLHLSLSLLANNAVTPNTPPPAQDSDHP 599
Db 527 SSLALYTMNSFFHGNISKEPLEINGL--SFGVDISPSNLQAEIRAGNAPLR-----576
QY 600 AVIGSTTAGSVTISGP--IFFEDLDDTA--YDRYDWLGSNQKINVLKQLGTGKPPANAPS 655
Db 577 -----LSGSPSIHDEGLFYENRDTAASPYQMEILLTSDKIVDISK-----TT 620
```

QY 656 DLTGEMPKYQGSWKLAWDPNTANNPY-TLKATWTKG-YNPGERVASLVPSLW 713  
DB 621 DSLVTNK--QSGFQGAWHFSPQNTINNTKQILRASWLPTEGYVLESNRVGRVAPVPSLW 678  
QY 714 GSILDIRSA-----HSAIQASVDGRSYCRGLWVSGV-----SNFFVHDDRDLGQ 757  
DB 679 STFLLLQTAHNLGDLHLCNNRSLIPTSYP-GVLIGTGAEMSTHSSEESFISRLGATGT 737  
QY 758 GYRIISGYSIGANSYFGSSMPLAFTEVFGRSKDYVVCRRNHHACIGSVY----- 808  
DB 738 SIIRLTPSLTLGG--GSHMFGDSF-----VADLPEHITSEGIQVNVGLTHVMG 784  
QY 809 -LSTQOALCGSYLFGDAPIRASYFGNQHMTSYTFABESDVWRDNNCLAGEIGAGLPV 867  
DB 785 PLTVNSTLCAA-LDHNAMVRIC-----SKOHTYG-----KWDTFQMRGTLGASYTFL 831  
QY 868 IPTSKLYNELRPPVQAESFYADHE-----SPTBEGDQARAPKSGHLLNLSVVPVGVKF 920  
DB 832 -----EYDQWTR-----VFSFANIEATNLIQRAFTETGYNPRSPSKTKLNLNIAPIGIGY 881  
QY 921 DRCSTHPNKYSFMA-----AYICDAYRTISGTETLLSHOBTWTD-----AFHLARHG 970  
DB 882 EFCLGN--SSPALLGKSGISYSDIKRENPSSTLAHLAMNDFAWTNGCSVPTSHTLANQ 939  
QY 971 VVVRGSMYASLTSNIEVGHGREYRDASRGYGLSAGSRVFP 1012  
DB 940 LILR---YKACSLYITAYTINR-EGKNLSN--SLSCGGYVGF 975

RESULT 13  
US-11-018-868-50  
; Sequence 50, Application US/11018868  
; Publication No. US20060034871A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; APPLICANT: Grandi, Guido  
; APPLICANT: Ratti, Gullio  
; APPLICANT: Bonci, Alessandro  
; APPLICANT: Finco, Oretta  
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis  
; FILE REFERENCE: 002441.00099 (PP23152.001)  
; CURRENT APPLICATION NUMBER: US/11/018.868  
; CURRENT FILING DATE: 2004-12-22  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 50  
; LENGTH: 975  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-11-018-868-50

Query Match 10.5%; Score 551; DB 11; Length 975;  
Best Local Similarity 24.9%; Pred. No. 2.4e-31;  
Matches 279; Conservative 148; Mismatches 405; Indels 290; Gaps 57;  
QY 3 TSPHKKFLLSMILAYSCSLNGGGVAAEIMV-----PQGIYDGETLTVFSPVTVIGDPSTG 57  
DB 32 TLIPKFLIGALIVVAPYSF-----AEMELAISGHKQG-KORDTFTM-----ISSCEPST 79  
QY 58 TVFSAGELTLK--NLDNSIAALPLSCFNLGLSPVTLGRGH--SLTPENIRTSNGAALS 113  
DB 80 NYIINRKLISDFSLNKKVSS--GGAFLNLAGKISFLGKNSASIFPKHININGFGAGVF 137  
QY 114 NSAADGLFTIRGKELSPNCSNLIALVLPAAATNKGSGOTPTTTPSPNGTIYSKTDLILL 173  
DB 138 SESS-----IEFTDLRLKLVAF-----GSES-----TCGIFTAKEDISFK 171  
QY 174 NNEKPSFYSNLVSGDGAIDAKSLTVQGISK-----LCVPQENTAOAGGACQVW 223  
DB 172 NNHIAFARNNTKNGGGVQI-----LQDMKGSVSFVDQGAIIPTNN-----QAV 217  
QY 224 TSPSAMANEAFIAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVSFSRNTAVEPDGN 283

DB 218 TS-SMKHSG-----RGCAISGDFAGR-----ILFLANQOITFSGN 253  
QY 284 VARYGGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT---SQQASNTSN---NYGDG 337  
DB 254 SAVHGGAIYN-----KNGLVEFLGN-AGP--LAPKENTTIANGGAIYTSNFRANQOTS 303  
QY 338 GAIFCKGQAQAGSNNSGSVSF-----DGBGVVFPSSNVAAKGKGAIAKLSVANCGP 390  
DB 304 PILFSQNH-----NKGCGAIYAQVNLNQODTIRFEKNTAKEGGGAISSQCSITAHNT 359  
QY 391 VOPLRLTAND--GCAIYL-GESGELSADVDGDIIPDGN--LKETAKEAADVNGVTYSS 445  
DB 360 IIFSDNAGDLGGGAILLEGKPSLTILAHSGNTAFSGNTWLHITTKASLDHRHSILIKR 419  
QY 446 QAISMGGKITTILKAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEYTT-----G 499  
DB 420 APYKI-----QLAANKNHSIHFFDPVMALSASSPIQ-----INAPETETPPSPKG 466  
QY 500 DIVP-----ANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAG 546  
DB 467 MIVFSGANLLDAREDEVANRTSIFNQPVHLYNGTSLIENGHAHLIVQSPKQTGGRISLSPG 526  
QY 547 STWDFVTPQ-----PPQOPPAANQLITLNLHLSSLILANNAVNTNPPTPPAQDSDHP 599  
DB 527 SSLALYTNWSPFHGNISKEPLEINGL--SPGVDISPENQARETRAGNAPLR----- 576  
QY 600 AVIGSTTAGSVTISGP--IPFEDLDDTA--YDRYDWLGSNQKINVLKQLGKTKPPANAP8 655  
DB 577 -----LSGSPSIHDPGLFVENRDTAASPVQMEILLTSDKIVDISKF-----TT 620  
QY 656 DLTGEMPKYQGSWKLAWDPNTANNPY-TLKATWTKG-YNPGERVASLVPSLW 713  
DB 621 DSLVTNK--QSGFQGAWHFSPQNTINNTKQILRASWLPTEGYVLESNRVGRVAPVPSLW 678  
QY 714 GSILDIRSA-----HSAIQASVDGRSYCRGLWVSGV-----SNFFVHDDRDLGQ 757  
DB 679 STFLLLQTAHNLGDLHLCNNRSLIPTSYP-GVLIGTGAEMSTHSSEESFISRLGATGT 737  
QY 758 GYRIISGYSIGANSYFGSSMPLAFTEVFGRSKDYVVCRRNHHACIGSVY----- 808  
DB 738 SIIRLTPSLTLGG--GSHMFGDSF-----VADLPEHITSEGIQVNVGLTHVMG 784  
QY 809 -LSTQOALCGSYLFGDAPIRASYFGNQHMTSYTFABESDVWRDNNCLAGEIGAGLPV 867  
DB 785 PLTVNSTLCAA-LDHNAMVRIC-----SKOHTYG-----KWDTFQMRGTLGASYTFL 831  
QY 868 IPTSKLYNELRPPVQAESFYADHE-----SPTBEGDQARAPKSGHLLNLSVVPVGVKF 920  
DB 832 -----EYDQWTR-----VFSFANIEATNLIQRAFTETGYNPRSPSKTKLNLNIAPIGIGY 881  
QY 921 DRCSTHPNKYSFMA-----AYICDAYRTISGTETLLSHOBTWTD-----AFHLARHG 970  
DB 882 EFCLGN--SSPALLGKSGISYSDIKRENPSSTLAHLAMNDFAWTNGCSVPTSHTLANQ 939  
QY 971 VVVRGSMYASLTSNIEVGHGREYRDASRGYGLSAGSRVFP 1012  
DB 940 LILR---YKACSLYITAYTINR-EGKNLSN--SLSCGGYVGF 975

RESULT 14  
US-11-013-957-11  
; Sequence 11, Application US/11103957  
; Publication No. US20050281847A1  
; GENERAL INFORMATION:  
; APPLICANT: Berthet, Francois-Xavier Jacques  
; APPLICANT: Lobet, Yves  
; APPLICANT: Poolman, Jan  
; APPLICANT: Verlant, Vincent Georges Christian Louis  
; TITLE OF INVENTION: Vaccine Composition  
; FILE REFERENCE: B45261  
; CURRENT APPLICATION NUMBER: US/11/103.957  
; CURRENT FILING DATE: 2005-04-12

```
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 013169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-11

Query Match          9.9%; Score 522.5; DB 11; Length 1034;
Best Local Similarity 23.6%; Pred. No. 2.9e-29;
Matches 274; Conservative 157; Mismatches 435; Indels 295; Gaps 53;

QY 14 LAYSCSLNGGGVAAEIMVPGIYDGETLTVSF-----PVTVIGDPSGTTVFSAGEL 65
DB 7 LSFACLSF---FYLSTISILQA---NETDTLQFRFTPSDRIFQVLDPA--SLITAQNI 58

QY 66 TLKNTLNSIAALPLSCFGLNLSFTVLGRGHSILTFENIRTSNGALSNSAADGL---FT 122
DB 59 VLSNLQS-----NGTGACTISGNTQTQIFSN---SVNTTADSGAFDMVTTSFT 104

QY 123 IEQFKELSFNSCNLSLAVLPAATNKGSGTPTTTSTPSNGTIYSKTDLLLLNNEKFSFYS 182
DB 105 ASDNANLLP--CNVY-----CTHNGK-----GGAIRSGGPIRFLANNQDVLFPN 145

QY 183 NLVSG-----DGAIDAKSLTVQGISKLVCVFOENTAQADGGACQVVTFSFA- 228
DB 146 NISAGAKYVGTGDHNEKNRGALYATTITLTG-NRTLAFINNMSGDCGGAISADTQISIT 204

QY 229 -----MANEAPIAFVANVAGGGAIAVQDGOQGVSSSTSTEDPVVFSRNT 276
DB 205 DTVKGLIFENNHTLNHPYTOAENMA--RGGAICSRD-----LCSISNNSGPVIV----- 252
QY 277 AVEFDGNVARVGGGIYSYGNVAFNNKTLFLNNAVSPVYIAAKQPTSGQASNTSNYGD 336
DB 253 ---FNTNQGKGGAISATRCVIDNKKERIIFSN-----SSLGMSSQSSASN 296

QY 337 GGAIFCRNGAQNNSGVSVPDGGVVPFPSSNVAAGKGAIYAKKLSVANCCGPQVPLRN 396
DB 297 GGAIQTTQGGFTL-RNNKGSYFD-----SNTATHAGGAINCGVIDIRKNDGVPVFLNN 347

QY 397 IANDGGAIVLGE--SGELSLSADYGDIIIP-----DGNLKRKTAKENAADVNGVTVSSQ 446
DB 348 SAANGAAFLNLSKPRSATNYIHTGTGDIVFNNVVFLDGNL--LGKELPHINNNEITPY 405

QY 447 AISMGSGGKITTLRAKAGHOILFND-----PIEMANGNPOAQQSSK-LLKINDGEGYTD 500
DB 406 TLSLG-----AKKDTRIYFDLPQWERVKENTSNNPSPSTRNTITVNPETEPFSGA 456

QY 501 IVFA-NGSSTLYQNV-----TTEQGRIVLRKAKLSVNSLSOTGGS---LYM 543
DB 457 VVFSYNQMSDRIITLMGKEHNYIKCAPTKLFGTLAIEDDAELEIFNPQTNPNTSLAL 516

QY 544 EAGSTWDFVTPPQQPQPAANQLITSLNLSLSSLLANNAVNTNP---TNPQAQSHPA 600
DB 517 GSGATLTV-----GRHGKLNITNLGVILPILKEG--KSPPCIRVNPQDMTQNTG 564

QY 601 VIGSTTAGSVTISGPIPF-----BDLDDTAYDRYDWLGSNOKINVLKQLQGTKPPAN 652
DB 565 T-GQTPSTSSISIPMIIFNGRLSIVDENYESVDSMDL--SRGKAQLILSIETTDGQ 621

QY 653 APSD--LTLGNEM---PKYGVQGSWKLAWDNT----- 680
DB 622 LDSNQSSLNTSLSPHYGYQGLWTPNWIITTYTITLNNSSAPTSATSIAEOKKTSET 681
QY 681 -----ANNG-----PYTLKATWTKTGYNPCGPVAVSLVPN 710
DB 682 FTPSNTTTTASIPNIKASAGSGSASNSGEVTTIKHTLVNVPVGYIVDPIRGDLIAN 741
```

```
Db 253 ---FNYNQGGKGAISATRCVIDNNKERIIFSN-----SSLGWSQSSASN 296
Qy 337 GGAIFCKNGAOGAGSNNSVSPDGGVFFSNVAAGKGAIYAKLSVANGCPVQFURN 396
Db 297 GGAIOQTQGTIL-RNNKGSIVPD-----SNTATHAGGAINCYDIDRNGPVYFLNN 347
Qy 397 IANDGAIYLCB--SGELSLADYGDIIIF-----DGNLKRTAKENAADVNGTVSSQ 446
Db 348 SAAMGAAPNLKSPRSATYIHTGTDIVNNNVFTLGNL--LCKRKLPHINNNEITPY 405
Qy 447 AISMGSGGKITTLRAKAGHQLFND-----PIEMANGNNQPAQSSK-LLKLDGSGYTD 500
Db 406 TISLG-----AKKDRIYFIDLPOWERKENTSNPPSPTSRNTITVNPTEFPGA 456
Qy 501 IVFA-NGSTLYQNV-----TIBQGRIVLREKAKLSVNSLSQTGS---LYM 543
Db 457 VVFSYNQMSDITRLMGKEHNYKEAPTTLKFGTLAIEDDAELEIFNIPFTQNPSTLLAL 516
Qy 544 EAGSTWDFVTPQPPQPPAANQLITLSNLHLSSLANNVTPNP-----THPPAQDSHPA 600
Db 517 GSGATLV-----GKHGLNITNLGVLPIILKEG--KSPPCIRVNPQDMQNTG 564
Qy 601 VIGSTTAGSVTSGPIF-----EDLDDTAYDRYDMLGNSQKINVLKQLGTPKPPAN 652
Db 565 T-GQTPSSTSIPTMIIIFNGELSIDENYSVDSMDL--SRGAEQLIISIEITNDGQ 621
Qy 653 APSD--LTLGNEM---PKYGGQSGKLAWDNP-----PKYGGQSGKLAWDNP 680
Db 622 LDSNQSSINTSLLSPHYGQGLWTPNWTITTTITLNNSSAPTSATSIAEQKKTSET 681
Qy 681 ---ANNG-----PYTLKATWTGTGYNPGPERVASLVPN 710
Db 682 FTSPNTTASIPNIKASAGSGSASNSGEVITTKHTLVVNAWPVGYTVDIRRGLIAN 741
Qy 711 SLWGSILDIRSAHSAIQASVDGRSICRGL-----WVS---GVSNEFFYHDDRDLG-QG 758
Db 742 SL-----VHS-----GRNWTGLRSLLPNSFWALQGAATLFTKQKRLSYHG 785
Qy 759 YRISGGYSLGANSYFGSMFGLAFTVFGRSKDYVCRSNHACIGSVYLSLTOALCGS 818
Db 786 YSSAKGYTVSQA---SGAHGKFLLSFGSSDKMKETNNRLSSRYLS---ALCFE 839
Qy 819 YLFGD--AFI-RASVGFQNHMKTSYTPAESDVDRWNNCLAGEIGAGLPVITPS-KLY 874
Db 840 HMFEDRIALIGAAACNYGTHNRSFYGKSGKGFHSTTL-----GASLRCELDRDSMPLR 895
Qy 875 LNELRPFVQAFSADHESFTEBGGQARAF--KSGHLNLNLSVPVGVKFDRCSSSTPN-KY 931
Db 896 SIMLTPFAQALFSRTEPASIRESGDLARLFTLEQHTAVS--PIGIKAYSDDTWPTLSW 954
Qy 932 SFMAAYICDARTISGTETLLSHOETWTTDAFLARHGUVVVRGSMYASLTSNIEVYHG 991
Db 955 EMELAYOPTLYMKRPLNTLLIQNGSNVTTNTPLAKHSFYGRGS-HSLKFSHLKLPANY 1013
Qy 992 RYERDASRGYGLSAGSRVP 1012
Db 1014 QAEVATSTSVHYINAGGALVP 1034
```

## RESULT 16

```
US-11-103-957-45
; Sequence 45, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
```

```
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1751
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-45
```

## Query Match 9.0%; Score 475.5; DB 11; Length 1751;

Best Local Similarity 23.6%; Pred. No. 1.5e-25;

Matches 269; Conservative 148; Mismatches 454; Indels 269; Gaps 51;

```
Qy 35 GIYDETLTVSPYTVVIGDPSTGTVFSAAGELTILKNLNSIAALPLSCFNNLLGSFTVLGR 94
Db 719 GIYAKKAMSRIDQLNISSENSATEI--GGGICCKESLELDALVSLVTENLVGK---BGG 773
Qy 95 GHSLTFFENIRTSWTGAALSNSAADGLFTIBGFKELSFSCNCSLLAVLPAATTNKGSOPT 154
Db 774 GLHAKTVNINSLKGFSPSNKANSSST-----GVATTASAPAAAASLQAAAAVPS 826
Qy 155 TTSTPSSNGTIYSKTDLLLNNEKFPYSNLVSGDGAIDAKSLTVQGISKLCVFQENTA- 213
Db 827 SPATPT-----YSGVV---GGAIYGEKVTFPSQCSGTCQFSGNQAI 863
Qy 214 -----QADGACQCVTVSFSMAANEAPAFV-----ANVAGVRGGGIAAVQD 254
Db 864 DNNPQSGLNVQGGAIYAKTSLSIGSDAGTSYFSGNSVSTGKSQTTGTQIAGG----- 917
Qy 255 GQQCVSSSTSTEDPVVFSRNTA-----VEPDGNVAR-----VGGIYSGNVAPLNN 302
Db 918 ---AIYSFTVTLNCPATFSNNTASMATPKTSSEDSGSGNSIKDTIGAI---AGTATLS 971
Qy 303 GKTLPLNNVA---SPVYIAAKQPTSGQASNTSNNYGD-----GGAIFCKNGA-QAGSNN 352
Db 972 GVSFRSGNTADLGAAGIAGTLANANTPSATSGNSITEKITLENGSPFIFERNOAKRGAIY 1031
Qy 353 SGSVSPDGGVGVVPSNVAAKGGAIIY-AKKLSVANCGPVQPL----- 394
Db 1032 SPSVSIKGNIT-FNQTSTHDSAIYFTKDATIESLGSVLFTGNNVTATQASATSQON 1090
Qy 395 RNIAWDGCAIYLGSGE-----LSLSADYGDIIIFDGNLKRKTAKENAADVNGTVSS 445
Db 1091 TMTANYGAALF-GDPGTTQSSQTDAILTLASSGNITFSNN---SLQNNQGD-----TPAS 1142
Qy 446 QAISMGSGGKITTLRAKAGHQLFNDPIEMAN---GNQPAQSSKKLKINDGEG---YTG 499
Db 1143 KFCISIAGVVXL-SLQAAGKGTISFFDCVHTSTKKIGSTQNVYET--LDINKENSNPYTG 1199
Qy 500 DIVPAN-----GSSTLYQNVITBQGRIVLREKAKLSVNSLSQTGS-LYMEAGSTWDFVTP 554
Db 1200 TIVFSELHENKSYIPQNAIHLNGTLVKEKTELHVVSFEQKSGSKLIMKPGAVL----- 1254
Qy 555 QPPQPPAANQLITLSNLHLSSLANNA--VTPN----- 590
Db 1255 ---SNQNIANGALVINGLTLDLSMWTPQAGEIFSPPELRIVATTSASGSGSVSSIPT 1311
Qy 591 NP-----PAQDHPAVITGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKINVLKQL 645
Db 1312 NPKRISAAAPSGSAATPTWSENKVFUTGLDLDLIDPNGNFYQN-PLMGSLDLVDVPLKIL-- 1368
Qy 646 GTEKPPANAPS-----DLTL-GNEMPKYQYQGSWKLAMPDNTANNPYYTLKATWTGTGNP- 699
Db 1369 ---PTNTSDVQVYDVLTLSGDLFPQKGYMGMTWTLDSNPQTK-----LQARWTFTYRRW 1419
Qy 700 -----GPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLGVSGV 744
Db 1420 VYIPRDNHFNANSILSGNSHIVVKQGLINNMLN-----NARFDDIAY-NNFWVSGV 1470
```

QY 745 SNFFYHDDALQGGYRISGGYSLGANSYFGSS--MFLAFTVFCRSKDYVVCRRNHAC 803  
 Db 1471 GTFLAQOQTPLSEESYYSRGTSVAIDAKPRQDFILGAAPSKWGTK--AIKWHNYFH 1528  
 QY 804 IGSVYLSQQALCGSYLFGDAFIR-----ASYGFGNQHMKTSY--TFAEES 847  
 Db 1529 KGSEYSYQASVYGGKFLY---PLLKQHGKVALPFLIQGVVSYGHKIKHDTTLLYPSIHERN 1585  
 QY 848 DVRWNNCLAGIEGAGLPVITTPSKLYNLNLPFVQAEFSYADHESFTEEGDQARAFKSG 907  
 Db 1586 KGDWEDLGLADLRISMDLK-EPSKSSKRITVYGELEYSSIRQKQFTEIDYDPRHFDCC 1644  
 QY 908 HLLNLSPVGVKFD-----RCSSTHNPYKYSFMAAYICDAVRTISGTETLLSHQETWTTDA 963  
 Db 1645 AYRNLSLPVGCAVEGAINMNCNLTMYNKLAL--AYMPSIYRNNPVCKYRVLSSNEAGQVIC 1702  
 QY 964 PHLARHGVVVRGSMYASLTNSI-----EYVGH-----GRYEYRDASRGYGLSAGSRVRF 1012  
 Db 1703 -----GVPTRTSARAESTQLYLGPFWTLGNYTIDVGMYTLSQMT-----SCGARMIF 1751  
 RESULT 17  
 US-11-018-868-16  
 ; Sequence 16, Application US/11018868  
 ; Publication No. US20060034871A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chiron Corporation  
 ; APPLICANT: Grandi, Guido  
 ; APPLICANT: Ratti, Giulio  
 ; APPLICANT: Bonci, Alessandro  
 ; APPLICANT: Fanco, Oretta  
 ; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis  
 ; FILE REFERENCE: 002441.00099 (PP23152.001)  
 ; CURRENT APPLICATION NUMBER: US/11/018,868  
 ; CURRENT FILING DATE: 2004-12-22  
 ; NUMBER OF SEQ ID NOS: 186  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 16  
 ; LENGTH: 1751  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia trachomatis  
 US-11-018-868-16  
 Query Match 9.0%; Score 475.5; DB 11; Length 1751;  
 Best Local Similarity 23.6%; Pred. No. 1.5e-25;  
 Matches: 269; Conservative 148; Mismatches 454; Indels 269; Gaps 51;  
 QY 35 GIYDGETLTVSPPTVIGDPSGTTVFSAGELTLKLDNSTAALPLSCFGLNLLGSFTVLGR 94  
 Db 719 GIYAKKAKMSRIDQLMISENSATEI--GGGICCKESLELDALVSLSTVENLVGK---EGG 773  
 QY 95 GHSLTPEINTRTSTNGAALNSAADGLFTTBGFKELSPNSCNSLLAVLPAATTNKGSTQPT 154  
 Db 774 GLHAKTWNISLXSGFSFNKANSST-----GVATTASAPAAAASLQAAARVPS 826  
 QY 155 TTSTPSTNGTIYSKTDLLLANNEKFSYNSLVSDGGAIDAKSITVQGISKLQVQENTA- 213  
 Db 827 SPATPT-----YSGVV---CGAIYGEKVTFSQCSGTCQFSGNQAI 863  
 QY 214 -----QADGACQVTVSPMANEAPIAV-----ANVAVRGGIIAAVQD 254  
 Db 864 DNNPSQSLNVQGAIIYAKTSLSIGSDAGTSYFSGNSVSTGKSQTTGQIAG----- 917  
 QY 255 GQGVSSSTSTEDPVVFSRNTA-----VBPQGNVAR-----VGGGIYSYGNVAFPLNN 302  
 Db 918 ----AIVSPVTLNCPATPSTNNTASMPATKTSSEDSGSGNSIKDTIGAI---AGTALTLS 971  
 QY 303 GKTLFLNNVA---SPVYIAAKQPTSGQASNTSNYGD-----GGAIFCKXGA-QAGSNN 352  
 Db 972 GVSRFSGNTADLGAAGTLANANTPSATSGNSITEKITLENGSPFIFERNQAMKRGAIY 1031  
 QY 353 SGVSVPDGEVGVFPFSSNVAACKGAIY-AKLSLVANCGPVQFL----- 394

Db 1032 SPSVSIKNNIT-FNQTSTHDGSAIYFTKDATIESGLSVLFTGNVNTATQASSATSON 1090  
 QY 395 RNIAIDGGAIVLGRSGE-----LSLSADYGDIIIPDGNLKRKTAKENAADVNGVTVSS 445  
 Db 1091 TMTANYGAALP-GPPTQSSQTDAILTLASSGNITFSNN---SLQNNQGD---TPAS 1142  
 QY 446 QAIMSGSGKITTLIRAKAGHOILFNDPIEMAN---GNNQPAOSSKLLKINDGEG---YTG 499  
 Db 1143 KFCSIAGVVKL-SLQAAAGKTIISFFDCVHTSTKKGSTQNVYET--LDINKENSENPYTG 1199  
 QY 500 DIVPAN-----GSSTLYQNWTTIEQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTP 554  
 Db 1200 TIVFSSSELHENKSYIPQAILHNGTLVLKERTELHVVSVFEQKESKLIKMPGAVL----- 1254  
 QY 555 QPPQPPAANQLITLSNLHLSLSSLLANNA--VTNP-----PT 590  
 Db 1255 ---SNQNTANGALVINGLUTIDLSMGTPQAGEISFPPELRIVATTSSASGGSGVSSIPT 1311  
 QY 591 NP-----PAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQNKINVLKQL 645  
 Db 1312 NPKRISAAAAPSGSRAATPTMTSENKVFLLGDLTLIDPNGNFYQN-PMLGSDLDVPLIKL-- 1368  
 QY 646 GTKPPANAPS-----DLTL-GNEMPKYIGQSGKWLAWDPNTANNPGPYTLKATTKTGYNP- 699  
 Db 1369 ----PTNTSDVQVYDLTSLGDLFPQKGYMGWTWLDNSNFTGK-----LQARWTFDTYRW 1419  
 QY 700 -----GPERVASLVNSLWGSITLDIRSAHSAIQASVDGRSYCRGLWVSGV 744  
 Db 1420 VYIPRDNHFIANSILGSONSMIVVQGLINMLN-----NARFDDIAY--NFWVSGV 1470  
 QY 745 SNFFYHDDALQGGYRISGGYSLGANSYFGSS--MFLAFTVFCRSKDYVVCRRNHAC 803  
 Db 1471 GTFLAQOQTPLSEESYYSRGTSVAIDAKPRQDFILGAAPSKWGTK--AIKWHNYFH 1528  
 QY 804 IGSVYLSQQALCGSYLFGDAFIR-----ASYGFGNQHMKTSY--TFAEES 847  
 Db 1529 KGSEYSYQASVYGGKFLY---PLLKQHGKVALPFLIQGVVSYGHKIKHDTTLLYPSIHERN 1585  
 QY 848 DVRWNNCLAGIEGAGLPVITTPSKLYNLNLPFVQAEFSYADHESFTEEGDQARAFKSG 907  
 Db 1586 KGDWEDLGLADLRISMDLK-EPSKSSKRITVYGELEYSSIRQKQFTEIDYDPRHFDCC 1644  
 QY 908 HLLNLSPVGVKFD-----RCSSTHNPYKYSFMAAYICDAVRTISGTETLLSHQETWTTDA 963  
 Db 1645 AYRNLSLPVGCAVEGAINMNCNLTMYNKLAL--AYMPSIYRNNPVCKYRVLSSNEAGQVIC 1702  
 QY 964 PHLARHGVVVRGSMYASLTNSI-----EYVGH-----GRYEYRDASRGYGLSAGSRVRF 1012  
 Db 1703 -----GVPTRTSARAESTQLYLGPFWTLGNYTIDVGMYTLSQMT-----SCGARMIF 1751  
 RESULT 18  
 US-11-103-957-21  
 ; Sequence 21, Application US/11103957  
 ; Publication No. US20050281847A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berthet, Francois-Xavier Jacques  
 ; APPLICANT: Lobet, Yves  
 ; APPLICANT: Poolman, Jan  
 ; APPLICANT: Verlant, Vincent Georges Christian Louis  
 ; TITLE OF INVENTION: Vaccine Composition  
 ; FILE REFERENCE: B45261  
 ; CURRENT APPLICATION NUMBER: US/11/103,957  
 ; CURRENT FILING DATE: 2005-04-12  
 ; PRIOR APPLICATION NUMBER: US/10/467,534  
 ; PRIOR FILING DATE: 2004-02-03  
 ; PRIOR APPLICATION NUMBER: PCT/EP02/01356  
 ; PRIOR FILING DATE: 2002-02-08  
 ; PRIOR APPLICATION NUMBER: GB 0103169.9  
 ; PRIOR FILING DATE: 2001-02-08  
 ; NUMBER OF SEQ ID NOS: 108  
 ; SOFTWARE: FastSeq for Windows Version 4.0





US-11-052-554A-1		Query Match		4.5%; Score 237.5; DB 11; Length 1268;	
		Best Local Similarity		20.8%; Pred. No. 1.4e-08;	
		Matches 257; Conservative 155; Mismatches 421; Indels 401; Gaps 66;			
QY	12	MILAYSCCSLNGGVAABIMVPOG-IYDGETLTVSPFVTVIGDPGTTVFSAGEITLKNL	70		
DB	21	MMLSPFPVAS-----QAAGLVKNGTVNAN---GVPVVDINKPNSGL-SHNIWDLNV	70		
QY	71	D-----NSTAULPLSCFGLGPTVLGRHSITPENIRTSSTGAALSNSAADGLFTI	123		
DB	71	DKNGVVFNNSANESSTSLAGNIQGN-SMLTSGSAKVIILNEVTSKPS-----TI	118		
QY	124	EGFKELSPSNCNSLIHLAVLPAATTKNGSQTP-----TSTSP-----SNGTI--	164		
DB	119	NGMEVAGDKADLIILANPENGITVNGGSGINTGKLTLTGTPTDIODDKLAGYVNGGTITL	178		
QY	165	-----YKTDLL-----LLNEKFPYSLVSGDGAIDAKSLTVQGISKLCYFQENTQA	215		
DB	179	GKLDNASPTILSRVVVNGKVSADELNVAGN-NYVNAAGQVTVGSVA-----TGS	230		
QY	216	DGACQVTVSPSAMANEAPIAFVANVAGVCGGIAAVODGQGVG-----	260		
DB	231	NGYSVDVAKLGMGYANKISLVSTEXGVGRNLGVIA--GGVNGVSDSKGNLLNSNAQIQ	288		
QY	261	-----SSTSTEDPVVSFS-----RNTAVEF-DGNVARVG-----	288		
DB	289	SASTINLTNGTLDNTTGTIVSGTISLNTNKTIVNTRAGNISITMGDIYVNSGTIDNTN	348		
QY	289	GGIYSGYGVAFNLNGKTLFLNNVSPV-----YIAAKOPTSGOASNTSNYDGGGAIFCK	343		
DB	349	GKLAAGMLAVDTNNTATLINSKSGSSVGTIAGLVALTKGT-----LNNSNGQIRGGVVGLE	404		
QY	344	NGAAGSNNSGVDFGEGVFPFSSNVAAGK-----GAIYAKKLKVANGC	389		
DB	405	SAAL--NNNGDIQTGTIAIISNGVNNKLIKRSSTGHIVIGAGSVNNGSTKTADTG	462		
QY	390	PVQPL-----RNANDGAIYLGESGELSLSA-----DYGDIIIFDGN---LKR	429		
DB	463	SDSLGIIADTQVEIGANNINNGGQI--ASNGVSLSSYSTIDDDYAGKILNSKVIKIG	520		
QY	430	TAKENAADVNGVTVSSQAISSMGSGKITTLRAKAGHQLFNDPIEMANGNPQASSKLL	489		
DB	521	SSLRN--DTGGIS-GKQIEVAVAGSLT-----NNIGVISSE--554			
QY	490	KINDGEGYTGD-I-VFAN-----GSSTLYQNVITIQ-----GRIVLEKAKLSV-NSL	534		
DB	555	-----EGDISLLANSVDNHHGFMGMQNIITMESMSGYNNNTALIVASKKLIKINARGSI	606		
QY	535	SQTGGSLYMEAGSTWDFVTPQPQPPAANQLITLSNLHLSLSLLANNA---VTNPPTN	591		
DB	607	ENRDGNF---GNAYGLYFGMPQ---TCGMVKGEGIELSGQNIYNNNSRLIAEDGPLT	659		
QY	592	PPAQ---DSHPAVIGSTTAGSVTISGPIFFEDLDDTDYDRY--DWLGSNQKINVLKQ--	644		
DB	660	LQANTFDNTRALVTSADASIQVGG-----TYNNYATTWSAGNLIDIDATTLQNS	710		
QY	645	-LGTKEPPANA-----PSDLTLG---NEMPKYGY--QGSWKLAWDPTANTANGP-YTLKAT	691		
DB	711	SSGTMIDNNATGFIASDKNLSEVNLSLTYNGWISGKGDVDF-----TVNNGNLNNTI	765		
QY	692	WTKYGNPQPERVASLVPNSLWGSIL--DI-RSAHSAIOASVDGRSYCRGLWVSGVSNF	747		
DB	766	AAEKGLD-----IAALNGIENKDISAGGDLTNTNTRHVTNNSNSNMVQNIIVINAVDI	820		
QY	748	-----PYHORD-----ALGQCYRYI-----SGGYSLGANSYFGSMFGLAFTVEFGSKDYV	794		
DB	821	NNRGNIVSDADLNVTTKGNLYLYVMVGYGDIALSANSVANNATTIATGDLIDSKGNV	880		
QY	795	-VCRSNHHACIGSV-----YLSQOALCGSYLFGDAPFASRYGFGNQHM	837		
DB	881	GNNRGNLHALNGVLVSKGNLNDNGEIRGVDVTLATGNY---DSY-----	925		

DB	1080	SSNGSQTDNLPLKLIASGNCIFRNNEYRPTSSDGTSTFCSIAGDVKLITQAAKGKTI	1139		
QY	469	FNDPIEMA-NGNQPAQSKLLKINDG-----GYTGDIVFAN-----GSTLYQNTYTI	516		
DB	1140	PFDAIRSTKTKGTQATAYDTLIDINKSBDSETVNSAFTGTILFSSSELHENKSYIPQNVVL	1199		
QY	517	EGGRIVLEKAKLSVNSLSQ--TGGSLYMEAGSTWDFVTPQPQPPAANQLITLSNLHLS	575		
DB	1200	HSGSLVLPKNTLHVLISFEQEGSSLVMTPGSVLSNOT-----VADGALVINNMVID	1251		
QY	576	LSSL-----LANNAVTNPPT-----NPPAQDSH-----	598		
DB	1252	LSSVEKNGIAEGNIFTPPELRIIDTTTGGSGGTPSTDSESNQNSDDTEONNNDASNOGE	1311		
QY	599	-----PAVIG-----STTAGSVTISGPIFFEDLDDTDYD	627		
DB	1312	SANGSSSPAVAAAHSTRTNPAATAATPTTTTATTTTNSQVILGGEIKLIDPNTFPQ	1371		
QY	628	RYDWLGSNOKINVLKLGTKPPANAPSLDTLGNEMPKYGYQSGWKLAWDPTANNPVT	687		
DB	1372	N-PALRSQOISLLVLPDSS--KMQAQKIVLTGDIAPQKGYGT--LTLDPOLQNG--T	1425		
QY	688	LKATWTKGYNPGERVASLVP-----NSLWGSILDIRSAHSAT-----QASVDGRS	734		
DB	1426	ISVLWKFDY-----QWAVYVRDNHIFYANSILGSQMLMVTYKQGLLNDKMLARFEVS	1480		
QY	735	YCRGLWVGVSFNFFHYDRDALOGYRYISGGYSLGANSYFGSSMP-GLAFTVEFGSKDY	793		
DB	1481	Y-NNLWISGLGTMLSGVQTPTEETYSRGSVALDAPKPAHDVIVGAAPFKMIGKTKS-	1538		
QY	794	VVCRSNHHACIGSVLSLTOALCGSVLFGDAPF-----ASYGFGNQHMKTS	840		
DB	1539	-LKRENNYTHKSEY-SYQASVYGGKPFHFVINKKTEKSLPLLQGVISYGIKHTVTH	1596		
QY	841	Y-TFAEESDVRDNNCLAGEIGA--GLPIVITPSKLYLNLRLPFPVQALPESYADHESFTEE	897		
DB	1597	YTIIRNKGEWED---LWLTLALRVSSVLRTPAQDGTKRITVYGELEYSSIRQKFTET	1653		
QY	898	GQOARAFKSHLLNLNVPGVVKPD-RCSSTHNKYS-FMAAYICDAYRTISGTETTLSSH	955		
DB	1654	EYDPRYFDCNTRNLAIPLMGLAPEGELSGNDILMNRFSVAYMLSIYRNSPTCKYQVLSS	1713		
QY	956	QETWTTDAFLAHGVVVGWSMYASLTSNI-----EVYCHGHEYERDASRGVLSAGSR	1009		
DB	1714	GEGGEIIC-----GVPTNRSARGEYSTQYLGLPLWTLYGSYTIADAHTLAHMMNCGAR	1767		
QY	1010	VRF 1012			
DB	1768	MTF 1770			

RESULT 20  
US-11-052-554A-1  
; Sequence 1, Application US/11052554A  
; Publication No. US2005028866A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 1268  
; TYPE: PR1  
; ORGANISM: Escherichia coli 0157:H7

```
Qy 838 KTSYTFEESDVRWNN-----CLAGE-----IGAGLPVITPSPKLYNELRPF 881
Db 926 KGSUT-SETGVTTTANIVDNAYGLIAGENVSVDASTIYNTALIAANKKLVIN----- 979
Qy 882 VQAFSVADHESFTEEGDQARAFSGHLLNLNSVPVGVKFDRCSSSTHPNKYSFMAAYICDA 941
Db 980 AGGNLENRDGNFLR-----NNGALPGITDVG----- 1007
Qy 942 YRTISGTETLLSHOETWTTDAFLARHG-----VVVRGSM----- 977
Db 1008 --GIVGKGVTLAQNVVNNSSIIAENGPLMLLSRGLDNRALLSGADAIIRAAGTF 1065
Qy 978 ---YASLTS--NIEVY-----GHGRYEYRDAS 999
Db 1066 YNNYATTYSAGNLDVYAAASLANNASDGRLEDNTAT 1099

RESULT 21
US-11-052-554A-172
; Sequence 172, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 172
; LENGTH: 1643
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
US-11-052-554A-172

Query Match 4.4%; Score 232.5; DB 11; Length 1643;
Best Local Similarity 19.6%; Pred. No. 4.7e-08;
Matches 218; Conservative 161; Mismatches 437; Indels 297; Gaps 53;

Qy 50 VIGDPSTVPSAGELTLKNLDNSTAALPLSCFGLNLSFTVLGRGHSHTPENIRTS--- 106
Db 678 VAADPLNTTTLDGTNLGSAENPLSTTHFATKAANADSILNVGKGVNLYANNITNDAN 737
Qy 107 -----TNGAALSNSAADGLFTIEGPKELSPS-----NCHSLLAVL 141
Db 738 VGSLLHFRSGGTSIVSGTVGGQGHKLNLLILDNGTTVKRFLGDTTFNGGKTIEGKSLQIS 797
Qy 142 PAATNKGSTPTTSTPSNGTI-YSKTDLLLLNNEKSPYSN-----LVSGDGAII--- 192
Db 798 NNYTTH-----VESADNTGLEFVNTDPTVTLNKGCAFGLVKQVILISGPNVINE 851
Qy 193 DAKSLITVQIGIKLCVQENTAAQ-----DGGACQVTSFSAWAM-----EAPIAFVAN 240
Db 852 IGVNIVHGIANSISFENASIGTSLPLSGTPLDLVLTIKTVCGNVTVDNFAPVIVVSG 911
Qy 241 VAGVRGG-----IAVQDQGVSSSTSTEDPVVPSRNT-----AVEFDGNVAVR 287
Db 912 IDSMINNQIIGDKKNIILSLGSD--NSITVNANTLYSGIRTTKNNQGTVTLSGGMENN 969
Qy 288 GGGIYSYGVNAFLNNGK-----TLFLNVASPVYIAAKQPTSGQASN--TSNNYGDG 338
Db 970 PTIYGLG-----LENGSPKLKQVTTTDDYNLGSII-----ANNVTIINY----- 1010
Qy 339 AIFCKNGAQAGSNNGSV---SFDGEGVVFFSSN-----VAAGKGAIIYAKKLS 384
Db 1011 -VTLTTGGIAGTDFDAKITLGSVNGNANVRVFDSTFSDPRSMIVATQANKGTVTYLGNAL 1069
```

```
Qy 385 VANGC-----PVQFLRNIANDGGAILYIGESGEL--SLSADYGDIIIPDGNLKRKAKENAADVN 439
Db 1070 VSNIGSLDTPVASVRFTGNDGA---GLQNIYISQNIIDFG-----TY 1108
Qy 440 GVTVSSQAISMGGGKITTLRAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTG 499
Db 1109 NLTLNSNVILGGG--TT-----AINGEIDL-----LTN 1135
Qy 500 DIVPANGSSLYQNVTIEQ-----GRIVLEBKAKLSVNSLSQTGGSLYMEAGSTWD 550
Db 1136 NLIPANGTSTWGDNTSISTTLNVSSGNIGQVIAEDAQVNAATTTTGTTIKIQDNANANF- 1194
Qy 551 FVTPQPPQPPAAQNLTLSNLHLSSLSSLLANNAVTPNPAPQADQHPAIVIGSTTAGSV 610
Db 1195 -----SGTOAYT-----LQGGARFNGTLCAP-----NFAVTGNS----- 1224
Qy 611 TISGPIF--FEDLDDTAYDRYDMLGSKNINVLKQLGTTPKANAP---SDLTIGNEMPK 665
Db 1225 -----IPVKYELIRDSNQD--YVLTRTNDVLNVVTTAVGNSAIAANAPGVSQNISRCLESTN 1278
Qy 666 YGQGSWKLAWDPNTANN--GPYTLKATWTKTGVN-PGPERSVASLVPSNLWGSI----- 716
Db 1279 TAAYNMMLAKDPDVAATFVCAIATDTSAAVTTVNLMDTKTQDLSNRL--GTLRLYLSNA 1337
Qy 717 --LDIRSAHSAIQASVDGRSYCRGLWVGVSVPFYHDDRDALGQGYRYISGYSISGANSYP 774
Db 1338 ETSDVAGSATGAVSSGDEAEVSYGWMAKPFYNIABQKGIAGYKAKTTGVVVGLDTLA 1397
Qy 775 GSSM-----FGLAFTEVFGRSKDYVVCVRSNHHCIGSVILS---TQALCGSLYFGDAFI 826
Db 1398 SDNLMICAAIGITKDI--KHQDY---KKGDKTDINGLSFSLYGSQQLVKNPFAQGNAPF 1452
Qy 827 RASYFGNQHMKTSTYTFABESDVR-----WNNCLAGEICAGLPIVITPCKLYNE 877
Db 1453 TL-----NKVSKSQRYEFESNGKMSQIAAGNYDNNMTFGNLIIFYDYNAMPNVL----- 1503
Qy 878 LRPFVQAEFSYADHESFTEGDDQA--RAFKSGHLLNLNSVPVGVKFDRCSSST-----H 927
Db 1504 VTPMAGLSYLKSSNENYKGTGTTVANKRINSKFSRDLVIVGAKV--AGSTVNTIDIVIY 1561
Qy 928 PNKYSFMA-----AYICDAYRTISGTETLLSHOETWTTDAFLARHGTVVVRGSMYASLTS 983
Db 1562 PEIHSFVHVHVKNGKLSNSQSLMDQQTAPFISQPDRTAKTSVNI-----GLSANIKS 1612
Qy 984 NIEV-YGHRYEYRDASRGYGLSAGS---RVRF 1012
Db 1613 DAKMEYGIG-YDFNSASK-YTAHQGTILKVRNF 1643

RESULT 22
US-11-052-554A-16
; Sequence 16, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 16
; LENGTH: 1250
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-16

Query Match 4.3%; Score 228; DB 11; Length 1250;
```

```

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1571
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-2

Query Match      4.1%; Score 218; DB 11; Length 1571;
Best Local Similarity 20.0%; Pred. No. 4.9e-07;
Matches 218; Conservative 138; Mismatches 413; Indels 320; Gaps 54;

QY 102 NIRTSTGAALNSAAGLFTIEGPKELSPNCNSLLAVLPAATNKGK-----QTPT 154
DB 102 NIRTSTGAALNSAAGLFTIEGPKELSPNCNSLLAVLPAATNKGK-----QTPT 154
QY 38 DIKASCHASRQSLGKITQWSTADQWLVSFD-----MTNNSAGGAVLQOGAE 86
DB 38 DIKASCHASRQSLGKITQWSTADQWLVSFD-----MTNNSAGGAVLQOGAE 86
QY 155 TTSTPSNGTIYSKTDLLLNNEKSPYSLVSGD---GGAIDAKSLTVQGISKLCVPOEN 211
DB 155 TTSTPSNGTIYSKTDLLLNNEKSPYSLVSGD---GGAIDAKSLTVQGISKLCVPOEN 211
QY 87 FTLSPEN-----ETGWL-----FANNVSGEYNNNGAIPAK-----EN 120
DB 87 FTLSPEN-----ETGWL-----FANNVSGEYNNNGAIPAK-----EN 120
QY 212 TAQADGGACQVVTSPSAMANEAPAFVANVAGRGGIAAVQDGGQGVSSSTSTEDPVVS 271
DB 212 TAQADGGACQVVTSPSAMANEAPAFVANVAGRGGIAAVQDGGQGVSSSTSTEDPVVS 271
QY 121 ST-----LNLTDVLFSGNVAGGYGAI-----YSSGNTDTCAL 153
DB 121 ST-----LNLTDVLFSGNVAGGYGAI-----YSSGNTDTCAL 153
QY 272 FSRNTAVBFDGNVAR--VGGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASN 329
DB 272 FSRNTAVBFDGNVAR--VGGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASN 329
QY 154 DLRTVNAFRNNIANDGKGAIYITNNDIYLSD--DVFNNOA-----YTS 197
DB 154 DLRTVNAFRNNIANDGKGAIYITNNDIYLSD--DVFNNOA-----YTS 197
QY 330 TSNVYDGGAIPECKGAQAGNSNSGVDFDCEGVV---PFSNVAAGKGAIAKLSVA 386
DB 330 TSNVYDGGAIPECKGAQAGNSNSGVDFDCEGVV---PFSNVAAGKGAIAKLSVA 386
QY 198 TSYSDGGAI-----DVTDNNSDSKHPSGYTIINNTAFTNNTAEGYGGAI-----T 245
DB 198 TSYSDGGAI-----DVTDNNSDSKHPSGYTIINNTAFTNNTAEGYGGAI-----T 245
QY 387 NCGPVQFLRNIAID-----GGAIYLGESELSADY 418
DB 387 NCGPVQFLRNIAID-----GGAIYLGESELSADY 418
QY 246 NSATAPYLIDISVDSYQNGVLVDENNSAAGYDGPSSAAGFPWYGLS-EVTFDIAD 304
DB 246 NSATAPYLIDISVDSYQNGVLVDENNSAAGYDGPSSAAGFPWYGLS-EVTFDIAD 304
QY 419 GDIIIDGNLKRTPAKENAADVNGTVSSQAIISMGSGGKITTLRAKAGHOIL-----FNDP 472
DB 419 GDIIIDGNLKRTPAKENAADVNGTVSSQAIISMGSGGKITTLRAKAGHOIL-----FNDP 472
QY 305 GKTLLVGN---TENDGAVD-----SIAGTGLIT--KTGSGDLVLNADNNDFTGE 348
DB 305 GKTLLVGN---TENDGAVD-----SIAGTGLIT--KTGSGDLVLNADNNDFTGE 348
QY 473 IEMANGNQPAGSSKLLKINDGE-----GYTCDIVPANGSSTLYQNVITIEQRIVL 523
DB 473 IEMANGNQPAGSSKLLKINDGE-----GYTCDIVPANGSSTLYQNVITIEQRIVL 523
QY 349 MQIENGVEVTLGRSNLSMLNVGDTHCQDDPDQCYGLT-----IGSIDKYQN-----392
DB 349 MQIENGVEVTLGRSNLSMLNVGDTHCQDDPDQCYGLT-----IGSIDKYQN-----392
QY 524 REKAKLSVNSLSQT-----GGSLYMEAGSTWDFVTPPPQPPQPPAANQLITLSNLHL 574
DB 524 REKAKLSVNSLSQT-----GGSLYMEAGSTWDFVTPPPQPPQPPAANQLITLSNLHL 574
QY 393 --QAEVLNUGSTQOTFAHSLTGFGQNTLNDAG-----422
DB 393 --QAEVLNUGSTQOTFAHSLTGFGQNTLNDAG-----422
QY 575 SLSSLLANNAVTPNPPAQSHPAVIGSTTAGSVTISGPIFFEDLDYDVRDVLGSG 634
DB 575 SLSSLLANNAVTPNPPAQSHPAVIGSTTAGSVTISGPIFFEDLDYDVRDVLGSG 634
QY 423 -----GNVTNQGSGFAGTIE-----GA 439
DB 423 -----GNVTNQGSGFAGTIE-----GA 439
QY 635 NQKINVKLQLGKTPPANAPSDLTLGNEMPKYQGSQWKLNDPNTANNPPTLKATWTK 694
DB 635 NQKINVKLQLGKTPPANAPSDLTLGNEMPKYQGSQWKLNDPNTANNPPTLKATWTK 694
QY 440 GQ-----LTIAQ-----NGSYVLAGAQSMAITGDIYVDAGAVL 472
DB 440 GQ-----LTIAQ-----NGSYVLAGAQSMAITGDIYVDAGAVL 472
QY 695 TGVNPGPERVA-----SLVPNSLMSIILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFF 748
DB 695 TGVNPGPERVA-----SLVPNSLMSIILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFF 748
QY 473 SLEGDAADLAALQDDPQSIVLN---GMLDL--SDPSTWQS---GTSYKQGLEVSGSGSTV 525
DB 473 SLEGDAADLAALQDDPQSIVLN---GMLDL--SDPSTWQS---GTSYKQGLEVSGSGSTV 525
QY 749 YHDRDALGQYRISGYSIGANSYFGSSMFLGAFTEVFGRSKD---YVVCRSNMHACIGS 806
DB 749 YHDRDALGQYRISGYSIGANSYFGSSMFLGAFTEVFGRSKD---YVVCRSNMHACIGS 806
QY 526 IGSQDVVD-----LAGNDM-----HIGDGKGQVYVVIDAGD---GQ 560
DB 526 IGSQDVVD-----LAGNDM-----HIGDGKGQVYVVIDAGD---GQ 560
QY 807 V-----YLSOQALCGSVLFDGAFIRASYGFGNQHKMTSYTPAESDV 849
DB 807 V-----YLSOQALCGSVLFDGAFIRASYGFGNQHKMTSYTPAESDV 849
QY 561 VSLANDNOVLGTQIASGTLMVSD---NSQLGYTHYNNRQVIPTDKPQESV 607
DB 561 VSLANDNOVLGTQIASGTLMVSD---NSQLGYTHYNNRQVIPTDKPQESV 607

```

RESULT 23

```

US-11-052-554A-2
; Sequence 2, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052, 554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20

```

```

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1571
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-2

Query Match      4.1%; Score 218; DB 11; Length 1571;
Best Local Similarity 20.0%; Pred. No. 4.9e-07;
Matches 218; Conservative 138; Mismatches 413; Indels 320; Gaps 54;

QY 21 LGGGYAAEIMVPOGIYDGETLTVSPFYTVLIGDPSGTVFVSAGELTLKLNDSIAALPLS 80
DB 21 LGGGYAAEIMVPOGIYDGETLTVSPFYTVLIGDPSGTVFVSAGELTLKLNDSIAALPLS 80
QY 416 INGG---TONINNYGIATGWINSGTQNIKSGKADTTIISSGSRQVVEKDTAIGSNIS 472
DB 416 INGG---TONINNYGIATGWINSGTQNIKSGKADTTIISSGSRQVVEKDTAIGSNIS 472
QY 81 CFGNLLGSFTVLGRGHSITFENIRTSNNGAALSNSAADGLFTIEGFKELS---FNCNSL 137
DB 81 CFGNLLGSFTVLGRGHSITFENIRTSNNGAALSNSAADGLFTIEGFKELS---FNCNSL 137
QY 473 AGGSLIVVTGGIAHG-----VNQETGSALVANTGAG--TDIEGYNKLSHPTITGGEAN 523
DB 473 AGGSLIVVTGGIAHG-----VNQETGSALVANTGAG--TDIEGYNKLSHPTITGGEAN 523
QY 138 LAVLPAATNKGSTPTTISTPSNGTI-----YSKTDLLLN-----EKFSFY 181
DB 138 LAVLPAATNKGSTPTTISTPSNGTI-----YSKTDLLLN-----EKFSFY 181
QY 524 YVVL-----ENTGELTVAKTSKNTIIDAGGKLIYQKEAKTDSTRLNNGVLLEVQDGBA 579
DB 524 YVVL-----ENTGELTVAKTSKNTIIDAGGKLIYQKEAKTDSTRLNNGVLLEVQDGBA 579
QY 182 SNLVSGDGAIDAKSLT---VQGISKL-CVF-----OENTAQADGGACQVVTFSAM-- 229
DB 182 SNLVSGDGAIDAKSLT---VQGISKL-CVF-----OENTAQADGGACQVVTFSAM-- 229
QY 580 KHVEQSGGALIASTTSGTLEGTSYGDAPYIRNSEAKNVVLENAGSLTVVTGSRVDT 639
DB 580 KHVEQSGGALIASTTSGTLEGTSYGDAPYIRNSEAKNVVLENAGSLTVVTGSRVDT 639
QY 230 ---AN-----EAPIAFVANVAG-----VRGG-----GIAA---VQDQGG 258
DB 230 ---AN-----EAPIAFVANVAG-----VRGG-----GIAA---VQDQGG 258
QY 640 IINANGKMDVYKDVGTVLNSAGTQTIYASATSDKANIKGKQTVYGLATEANIESGEQI 699
DB 640 IINANGKMDVYKDVGTVLNSAGTQTIYASATSDKANIKGKQTVYGLATEANIESGEQI 699
QY 259 VSSSTSTEDPVVPSRNTAVEF-----DGNVARVGGGIYSYGNV 297
DB 259 VSSSTSTEDPVVPSRNTAVEF-----DGNVARVGGGIYSYGNV 297
QY 700 VDGG--STERKTHINGTQTVQNYKCAINTDIVSGLQOIWANGTAEGSIINGGSI VNEGGL 758
DB 700 VDGG--STERKTHINGTQTVQNYKCAINTDIVSGLQOIWANGTAEGSIINGGSI VNEGGL 758
QY 298 A-----FLNNGKTLFLNNVASPVYIAAKQPTSG-----QASNTSNNGDGGAI FCKNGAQA 348
DB 298 A-----FLNNGKTLFLNNVASPVYIAAKQPTSG-----QASNTSNNGDGGAI FCKNGAQA 348
QY 759 AENSVLNDGGTLVDREKGSATGI--QSSQSGALVATTATRTVGTGTRADGVAFSIQQGA-- 814
DB 759 AENSVLNDGGTLVDREKGSATGI--QSSQSGALVATTATRTVGTGTRADGVAFSIQQGA-- 814
QY 349 GSNNSGSVSFDGEGVYVFFSNVAAGK---GGAIYAKKLSVAN---CGPVQFLRNIA- 398
DB 349 GSNNSGSVSFDGEGVYVFFSNVAAGK---GGAIYAKKLSVAN---CGPVQFLRNIA- 398
QY 815 -ANN---ILLANGVLTVESDTSKQTVNTGGRBIVKTKATATGTTLTGGGQIVEGVAN 870
DB 815 -ANN---ILLANGVLTVESDTSKQTVNTGGRBIVKTKATATGTTLTGGGQIVEGVAN 870
QY 399 ---NDGGAIVLGRSGELSLSA--DYGDIIIFGNIKRT-----AKENAADVNGTVSS 445
DB 399 ---NDGGAIVLGRSGELSLSA--DYGDIIIFGNIKRT-----AKENAADVNGTVSS 445
QY 871 ETTINDGGIQTVSANGAEIKTTINEGGTLTVNDNGKATDIVQNSGAALQTSTANGIEISG 930
DB 871 ETTINDGGIQTVSANGAEIKTTINEGGTLTVNDNGKATDIVQNSGAALQTSTANGIEISG 930
QY 446 Q---AISMGSGKITTLRAKAGHOILF-----NDPIEMANGNQPAGSSKLLKINDGE 495
DB 446 Q---AISMGSGKITTLRAKAGHOILF-----NDPIEMANGNQPAGSSKLLKINDGE 495
QY 931 THQYGTFTSI--SGNLATNMLLENGNLLVLAGTEARDSTVKGGMAMQOQDSATKNSGG 989
DB 931 THQYGTFTSI--SGNLATNMLLENGNLLVLAGTEARDSTVKGGMAMQOQDSATKNSGG 989
QY 496 GYT-----GDIVFANGSSTLY-----QNVITIE-- 517
DB 496 GYT-----GDIVFANGSSTLY-----QNVITIE-- 517
QY 990 QYTLGRSKDEFOALARAEDLOVAGGTAVIYAGTLADASVSGATGSLATPRDNVTPVKL 1049
DB 990 QYTLGRSKDEFOALARAEDLOVAGGTAVIYAGTLADASVSGATGSLATPRDNVTPVKL 1049
QY 518 QGRIVLREKAKLSV--NSLSQTGGSLYMEA--GSTWDFVTVPQPPQPPAANQLITLSNLHLS 575
DB 518 QGRIVLREKAKLSV--NSLSQTGGSLYMEA--GSTWDFVTVPQPPQPPAANQLITLSNLHLS 575
QY 1050 EGAIKITDSATLTIGNGVDTTLADLTAAASRGVW-----LNSNNSCAGTSNCEYR 1099
DB 1050 EGAIKITDSATLTIGNGVDTTLADLTAAASRGVW-----LNSNNSCAGTSNCEYR 1099
QY 576 LSSILLAN--NAVNPPTNPPAQDSDHPAVIGSTTAGSVTISGPIFFEDLDYDVRDVLG 633
DB 576 LSSILLAN--NAVNPPTNPPAQDSDHPAVIGSTTAGSVTISGPIFFEDLDYDVRDVLG 633
QY 1100 VNSLLNDGNVYLSAQTAAPATTN--GIYNTLTNELSGSGNFYLH--TNVAGSRGDLV 1155
DB 1100 VNSLLNDGNVYLSAQTAAPATTN--GIYNTLTNELSGSGNFYLH--TNVAGSRGDLV 1155
QY 634 SNQ-----KINVKLQLGKTPKAN-----APSDLTLGN-----EMPKYGY----- 668
DB 634 SNQ-----KINVKLQLGKTPKAN-----APSDLTLGN-----EMPKYGY----- 668
QY 1156 VNNNATGNFKI FQDTGTVGSPQSDAMTLVKTGGDASFSLGNTGGFVDTLGTVEYVLKSDG 1215
DB 1156 VNNNATGNFKI FQDTGTVGSPQSDAMTLVKTGGDASFSLGNTGGFVDTLGTVEYVLKSDG 1215
QY 669 QGSWKLAWD--PN--TANNGPYTLKATWTKTVGNYPGE-----702
DB 669 QGSWKLAWD--PN--TANNGPYTLKATWTKTVGNYPGE-----702
QY 1216 NSNNLTNDVKPNPDNPNPNPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDP 1275
DB 1216 NSNNLTNDVKPNPDNPNPNPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDPKDP 1275

```

Qy 703 --RVASLVP-----NSLWGSILDIRS-----AHSATQASVDG 732  
Db 1276 VLNMAATLPLVFDIAELNSIRERLNMKASPINNNVWGATYTRNNVTVDAGAEQTLTG 1335  
Qy 733 -----RSYCRGLWVGVSFFVHRRDALGQGYRISYGGYSLGA-NSYFGSSMFLGAF 783  
Db 1336 MTVGIDSPNDIPEGIATLGAFWGSHSHIGDRGHSVGYSYLGVSASHEHSGFYLDG 1395  
Qy 784 TEVGRSKDYVVCRRNHACIGSVYLTQOALCGSYLFGDAFIRASYGFCRQHMKTSYTF 843  
Db 1396 VVKLNRPSNVAGK-----MSSGGAANGSY-----HSNGLGG-HIETGHRF 1435  
Qy 844 AERSDVRWD 852  
Db 1436 ---TDGNWN 1441

RESULT 24  
US-11-087-099-1245  
; Sequence 1245, Application US/11087099  
; Publication No. US20060041961A1  
; GENERAL INFORMATION:  
; APPLICANT: Abad, Mark S. et al.  
; TITLE OF INVENTION: Genes and Uses for Plant Improvement  
; FILE REFERENCE: 38-21(53450)B EP  
; CURRENT APPLICATION NUMBER: US/11/087,099  
; CURRENT FILING DATE: 2005-03-22  
; NUMBER OF SEQ ID NOS: 12464  
; SEQ ID NO 1245  
; LENGTH: 3132  
; TYPE: PRT  
; ORGANISM: Magnetococcus sp. MC-1  
US-11-087-099-1245

Query Match 4.1%; Score 218; DB 11; Length 3132;  
Best Local Similarity 20.6%; Pred. No. 1.3e-06;  
Matches 228; Conservative 144; Mismatches 434; Indels 304; Gaps 51;

Qy 23 GGGYAAEIMVPGGIYD-----GETLVSPFYTVIGDPSGTTV-FSAGELTLKN----- 69  
Db 1035 GGGHVDVINAISDGVLSLGMVSAVGSVTIDGGTGSILDNNGATANIANSASIKSDKAG 1094

Qy 70 -----LNSIAALPLSCFNLGSEFTVLGRGHSITFENIRT-----STN-----GAALSNSAA 117  
Db 1095 TASDALETTVDTLAVEITGS--GKSFYIDESDALTSINAKVNGSTNLNFTGGSFAFNAT 1152

Qy 118 DGLFTIEGFKELSPSNCNSLLAVLPAATTNKGSTPT-----T 155  
Db 1153 TGAFSSTGVGVTFENTGGVVAI-GTVTATGGSATITATTAITATSATADTVVLNAGT 1211

Qy 156 TSTPSNGTIYKTDLLLNNEKFSFYSLNVLGSDGDAIDAKSLTVQGISKLCVFOENTAQA 215  
Db 1212 SIGASGSTIKSTATTLTLVASAGSIYAQ--BSDGATVNKAV--GTGKNIEFATTGN- 1265

Qy 216 DGGACQVTSFSAMANEAPAFVANVAGVGGGIAVQDQGGVSSSTSTEDPVVPSRN 275  
Db 1266 -----LTGTTISAKGSVTLTATAGSVLSGG-----TSSSATGATVLSAG 1305

Qy 276 TAVRFDG-----NVARVGGIY--SYGNVAF-----LNNKTLFLNNVA 312  
Db 1306 TAIGASGASVTVVAANLSATASAGVYLSNAGDVTLTAVATGAGFQLSNTGALVNLGVT 1365

Qy 313 SPVYIAAKQPTSGQASNTSNNGYDGGAIKCKNGAQAGSNNSGVSFDCGEGVVFSSNVAA 372  
Db 1366 -----AAQAVSLTAS-----GALTDCNGA--TNNISABSLTLVGLSIGST----- 1404

Qy 373 KGGAIIYAKKLSVANGCPVQLRNANDGGAIYIGE--SGELSLADYGDIIIFQNLKRT 430  
Db 1405 ATNGAVDTQVNSV-----TATTTSGGIIYNELSGGLTITAATA-VGSDANVSIT 1453

Qy 431 AKENAADVGVTVSSQATISMGSGGKITTLRAKAGH-----QIL 468

Db 1454 GAGDIA-LGVITAKGDDVTLVSAGKITDNNGANNVTTADILNVTGPNVGELETISITQLS 1512  
Qy 469 FNDPIEMANGNPOAQSGL-----LKINDGEGYTGDIIVFANGSSTLYQNVVIE 517  
Db 1513 STGSTDIVNAGAMAITKASLEGSSSPFAESLTLDMAGDTATI--ANNISLTQLTTT-- 1568  
Qy 518 QGRIVLEKAKLSVNS--LSQTGSLYMEAGSTWDFTPQBPQPPAANQLITLSNLHLS 575  
Db 1569 -GNIVF-----LQNDTIVAQSGTGTINAGTT-----DKGAVAIIGNITANONIS 1615  
Qy 576 LSSLLANNAVTPPTNPADSHPAVIGSTTA--GSVTISGPIFFEDLDDTAYDRYDWLG 633  
Db 1616 ITA-----DSH-ITIGLNLNAGIDGVSVD-YGVLDD-----G 1646

Qy 634 SNOKINYL--KLQLGTKPPANAPSDLTLGNEMPK---YGYGGSWKLA-WDPNTA----- 681  
Db 1647 NSTAVNVIAARNFSLSGTTPTRQAEHLHTTNSIANAHADSEVAAKLTLLLEANTAAMDIMS 1706

Qy 682 ---NNGPYTLKATWTKTGYNPGERVASLVPNS-----LWGSILDIRSAHSAIQAS 729  
Db 1707 TABSTANYSL--TLASSAYDTAQAEVDRLAPIENGLKITYVLDGVSIALSTAASAIEVA 1764

Qy 730 VDCRSYCRGLWVGVSNNFFYHRRDALCQGYRYISGGYSLGAN-SYFGSSMFLGAFTEVFG 788  
Db 1765 -----AGAAQAIPPSGDAGSEAAAVTG---LAANVAGIAAAYAIGIAHSEIAG 1809

Qy 789 RSKDYVVVCRNHHACIGSVYLSLTOALCGSYLFGDA--FIRASYGFCRQHMKTSYTFABE 846  
Db 1810 QLDDAEDLAFNKDAEMVAAKSTLTDLAIHQAYKEATSISQAAVDAAVIERHNHMQVARQ 1869

Qy 847 SDVRWNNCLAGEIGAGLPVITPISKLYLNELPFVQAEFSYADHESFTBEGDQARAFKS 906  
Db 1870 AVTAEDASSTAGTLALPFGIQIS-GQLDVNAV-----NSDVVLAITG 1910

Qy 907 GHILAN-LSVPVGKFCRCSSTHPNKYSFMAAYICDAVRTISGTETLLSHOETWTTDAPH 965  
Db 1911 PAVLNSASATKTAGTGMISITATDNISVQALVADDIURL---ETT----- 1953

Qy 966 LARHGVVVVGRSMYASLTSN--IEVYGHG 991  
Db 1954 ---GQIINGNAGHSLTASKFVGVA GTG 1977

RESULT 25  
US-11-052-554A-90  
; Sequence 90, Application US/11052554A  
; Publication No. US20050288866A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; FILE OF INVENTION: PROTEINS OF THERAPUTIC POTENTIAL  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 90  
; LENGTH: 3194  
; TYPE: PRT  
; ORGANISM: Helicobacter pylori J99  
US-11-052-554A-90

Query Match 4.1%; Score 218; DB 11; Length 3194;  
Best Local Similarity 19.8%; Pred. No. 1.4e-06;  
Matches 229; Conservative 126; Mismatches 403; Indels 396; Gaps 56;

Qy 14 LAYSCCSLNGGGYAAETMWPGGIYDGETLATVSPFYTVIGDPSGTTVFSAGELTLK----- 68  
Db 231 ITYNGSVNGNFGFDNVDNGA-----TTI---SGVTFNNNGALTYKGGNGI 275

QY	69	-----NLDNSTAALPLSCFG-----NLLGSPVTLGRGHSLTPEINRT	105
DB	276	GGSIITFTNSINHYKLNINANSVTFNNSTLGSMPNGNANTIGNAYIL--NANNITFNNL--	332
QY	106	STNGAALSNAADGLFTIEGPKELS-----FSCNSSLAVLPAAATNKGSTOPT	154
DB	333	TFNGGWFVFNRSDAHVNFGQTTINPTSPFVNMTGKVTINPAINFNQNTYPTIGNAYT	392
QY	155	-----TTSTPSN-----GTIYS---	166
DB	393	LFSMKGNIAYDVNVLNWIIRLKNTOATKNSKNATSNNTHTYVTVYVTLGGTILYHRQ	452
QY	167	-----KTDLLLLNNEKFS-----FYSNLSVGGGADAKSLT-----VOGISKLCVFOEN	211
DB	453	IFSPDSIVLQSVYYGANNLYTNSVNIHNVFNLNKINDDRADTIFYLNGLMTWNTQAR	512
QY	212	TAQADGACQVVTSTFSAMANEAPIAFVANAVGRGGIAAVQDGGQGVSSSTSTEDPVVS	271
DB	513	FAQTGGKNSALV-FNAITPWANGALPKNSVTRFGYEGVWNGKTYITGFTADRVI	571
QY	272	FSRNTAVEPDGNVARVGGG-IYSYGNVAFNLNKGTLF--LNNVASPVYIAAKOPTSGQAS	328
DB	572	-----IGNMMSGNAQTGGCATLNFVCGATEINIAGATFKNLKTTSQNSYMTFMALNGSGS	627
QY	329	NTSN-----NYDGGAIFCNKGQAQGSN-NSGSVSPDG--EGVVPFSSNVAAG-----K	374
DB	628	GKINVSQSDFYDWTDCGYDFTGNGVFDVSVPFNKAYYKFQGAENSYNFKNTNFLAGNFKQ	687
QY	375	GCAIYAKKLSVANGCPVQF--LRNIAND-----GGAIYLGES-----GELSUS	415
DB	688	GKTTIEK--SVLNDASVAFDGVNNAFNEKFGNGSFNFHAEQTNAPNNSFGSGFSFN	745
QY	416	A---DYGDIIIFDQ-----NLKRTAK-----ENAAVDNGVTVS---SQAISMSGSGGKI	456
DB	746	AKQVDNGNSFNGGVFNENFTPKASFTNDTFVNNQFKINGAQDTPEKGVVFNMOGLL	805
QY	457	TTLRKAGHQIILPNDPIEMANGNQPAQSKLLKINDGEGYTGDIY-----	502
DB	806	SSLSVGTYYQLLNAKSVGYKDNNNALYQ---MLRWTSGENPSGKLVDEKNTAPNSAKIYN	862
QY	503	--PANGSSTLYQNVTIEQRIVLREKALSV-----NSLSQTGSSLYM	543
DB	863	VQPTDNGLYIYIKENFNNG-ITUTRLCTLYHCVIDNDADFNLKVNNNASNT--VFYL	919
QY	544	EAGSTWDFVTPQPPQPPAANQILITLSNLHLSLSLLANNAV-----NPPTNPPAQDS	597
DB	920	NGMTW-----KTAGTVFTQD--YSGTNSVLVFNQTPPLAGANPTSN-----	961
QY	598	HPAVIG-STTAGSV-----TISGPPIFEDLDDTAYDRYDVLGWNQKINVLKQLGTYKPPA	651
DB	962	--SVWFGKTSAGWGLVGYIQGVFKANQIDIT-----GTIRSGNAKTCGGATLVF	1011
QY	652	NAPSDLTLYN---EMPYKYGQSKLAWDPNTAN-----NGPYTLKA---TW	692
DB	1012	NAQERLIANANLNDKAGLQNSW-MNFIVNNGNLNVNANFNQTPHGFNFKANNITW	1070
QY	693	TKTGYNPGBPERVASLVNSLWGSILDIRSAHSAIQASVDRSVCRLGWSGVSNFFYHDR	752
DB	1071	DK-----GS-----VSGGNGFVDNA	1086
QY	753	DALG-----QGYRYISGG---YSLGANSYFGSSMFLAFTVFGKSDYVVCRRNHACIG	805
DB	1087	NANGNAVIMKNVFNSDNCTLYKGGENS-----AGNSLTLENNTFNSY-----NINAKA	1135
QY	806	SVYLSSTQALCGSVLFCD-----AFIRASVFGNQHM-----K	838
DB	1136	NLIFNNNSFNSGSYSPFNDTKNVTFKGTNTILNSDPFSLKGSVSDINNSIFNIERLDTDK	1195
QY	839	TSYTFABESDVRWDDNCLN-----GEIGAGLPIVITPSKLYLNLNLRPFVQ	883
DB	1196	TTTTLLSGDNIKYNNOALADNVFSKNLWDLIHVDEGEQT---LLRTDNTY-----FVQ	1246

QY	884	AEFSYADHESFTEE	897
DB	1247	--FTQSNQKQKVPB	1258

RESULT 26

US-11-052-554A-281

Sequence 281, Application US/11052554A

Publication No. US20050288866A1

GENERAL INFORMATION:

APPLICANT: Sachdeva, et al.

TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE

FILE REFERENCE: 30853/40359A

CURRENT APPLICATION NUMBER: US/11/052.554A

CURRENT FILING DATE: 2005-02-07

PRIOR FILING DATE: 2004-07-20

PRIOR APPLICATION NUMBER: IN 173/DEL/2004

PRIOR FILING DATE: 2004-02-06

NUMBER OF SEQ ID NOS: 763

SOFTWARE: PatentIn version 3.3

SEQ ID NO 281

LENGTH: 5291

TYPE: PR

ORGANISM: Escherichia coli 0157:H7

US-11-052-554A-281

Query Match 4.0%; Score 213; DB 11; Length 5291;

Best Local Similarity 21.8%; Pred. No. 6.5e-06;

Matches 182; Conservative 116; Mismatches 360; Indels 176; Gaps 33;

QY	39	GETLV-----SPPTVIGDPSGTTVPFSAAGELT-LKLNDSIAALPLSCFGLN-----	85
DB	3267	QGLTTLVTLNGTYQTTVTDGWSLTLTPASDALTALANNGYTATVSDLAGLMSKSGV	3326
QY	86	-----LGSPTVLGRGHSLTFFENIRTS--TNGAALSNSAADGL-FTIEGFKELSFSCN	135
DB	3327	TYDTTAPVISEFTVAGDDVINNVVHIQAQIISGTATGAVAGDLVVTTIAGQQVYVTS	3386
QY	136	SLLAV-LPAATN---KGSQPTPTTTPSNGHYISKDLLLLNNEKFSYSLVSGDG--	189
DB	3387	GNWSGVGPASVISGLADGTVTISATITDSAGNSSTQTHNVQNTAAVLSVSTISGDNLI	3446
QY	190	GAIDAKS-LTVQGI-SKLCVFOENTAQADG-CACQVVTFSFSAWNEAPLAFVANVA-GVR	245
DB	3447	NAABAGSALTLSGTGNFATGVTVVLLNGKGYSATIQSGWSVNVNPAADVAALSDGTS	3506
QY	246	GGGIAAVDQCGQCVSSSTSTEDPVVSFSRNTAVFEDGNVARGGVIYSGNVAFNLNGKT	305
DB	3507	YTVSASAQDSAGNSSTQTHNVQV-----NTAAVLSVSTISGDNLIN---3550	
QY	306	LFLNNVASPVYIAAKOPTSGQASNTSNNYDGGAI-FCRQG-----AQAGNSNSGVS	360
DB	3551	-----AABAGSALTLSGTGTNPFATGVTVVLLNGKGYSATIQSGWSVNVNPA--	3597
QY	361	EGVVPFSSNVAAGKGAIIYAKKLSVANCCPVQFLRNIANDGGAIIYLGESGELSLSA--	417
DB	3598	-----AADVAALSDGTSYTVSASAQD-----SAGNSATASRVAVDLTAPVTS	3640
QY	418	YGDIIIFDGNLRTAKENAADVNGVTVS--SQAISMSGSGGKIITTLRKAQHILFNDPIEM	475
DB	3641	INTVSTDDRLNAAEQOPLTLNGSTSAEVQTVTVTFGKTYTATVAANGTALNVPADV	3700
QY	476	ANGNPPAQSSKLLKINDGEGYTGDIYFANGSSTLYQNVTIE--QGRIVLREKAKLSVNS	533
DB	3701	LAALGQGAQTI-TASVNDRAGNPGQATHALTVDVTAPTVTIATVAGDDIINNAEQLAGQT	3759
QY	534	LSQT-----GGSLYMEAGSTWDFVTPQPPQPPAANQILITLSNLHLSLS	577
DB	3760	ISGTTTTAEVQGTVTVTFNGQTSATVSGSGSWSVFIP-----AQQFAGLSDDGYTIS	3811
QY	578	SLLANNAVTPPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDVLGWSNQK	637





382 RTNPIQACTSPSAPITVTGTAALL--QKYPWMSN-----DNLRTT 420  
QY 209 QENTAQADGACQOVVTSF-----SAMANEPIAPVANAVRGGGIAAQQDGGQVYS 260  
Db 421 LLTTAQ-DIGAVGVDSKFGWLLDAGKAMNGPASFPFGDTADTKG----- 465  
QY 261 SSTSTEDPVVPSRNTAVFPGNVARVGGGITYSGNVAF-----LNNKGTFLNNVASPV 315  
Db 466 -----TSDIAYSP-RNDISGTGGLIKGGSQLQHGNNTYTKTIIIEGSSLVLYGNKSDM 520  
QY 316 YIAAKQPT--SGQASNTSNNGDGAIFCKNGAQAGSNNS-----GVSFDFGEGVFFSSN 369  
Db 521 RVETKALYNGAAGSGSLN--SDGIVYLADTQSGANETVHIKGSLLQDGGKTLX----- 574  
QY 370 VAAGK-----GGAIYAKK-----LSVANGC-PVQFLRNITANDG 401  
Db 575 TRLGKLLKVDGTAIIGGKLYMSARGKAGYLNSTGRRVPFLSAAKIGQDYSFFTNITD 634  
QY 402 GAIYLGESGELSADYGD-----IIPDGNLKRITAK-----ENAADVNGVTVSSQA 447  
Db 635 GLLASLDSVEKTAGSE-GDTLSYVVRGNAARTASAAHSAAPAGLKHAVEQGSNLENLM 693  
QY 448 ISMG-----SGKITTLLAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFA 504  
Db 694 VELDASESSATPETVETAA-----DRTDMPGIRPVYGFATRAAAAVQHANAADGVRIEN 747  
QY 505 NGSTLYQNVITB-----QGRIVUREKAKLSVNSLSQTCGSLYMEA-----GSTWD----- 550  
Db 748 SLAATVYVDAAHADMQGR-----RLKAVSDGLDHNGTGLRVIAOTQDDGGTWEQGV 802  
QY 551 -----FVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPPAQDSHPA 600  
Db 803 GWRGSGTQTVGIAAKGENTAA-----ATLGMGRSTWSENSE--NAKT-----DSISL 849  
QY 601 VIG-STTAGSVTISGPFFEDLDDTAYDRYD-----WLGSNQ-----KINVLKQLGKTKPP 650  
Db 850 FAGIRHDAGDIGYKGLF-----SYGRYKNSISRSTGADAEHAGSVNGTLQGLALGG 902  
QY 651 ANAP-----SOLITGN-----EMPKYGYOGSKLAWDPYIANG-----PYT 687  
Db 903 VNVFPAATGDLTVEGGLRYDLKQDAPAEKGS-ALGWSGNSLTEGTLVGLAGLKLQPLS 961  
QY 688 LKATWTGTGYNPGERVASLPVNSLWGSILLDIRSA-----HSAIOASVDGRSYCR 737  
Db 962 DKAVLFATA-----GVERDLNGRDYTVTGGTGATAATCKTGARNMPHRLVAGLADVEFG 1018  
QY 738 GLWVSGVSNFFY-----HSDRALGQGYRYISGYSILG-----ANSYFGSSMFLAFTV 786  
Db 1019 NGW-NGLARYSYAGSKQYGNHSGRVGVYRFLDGGGCTGSSDLANDSFIRQVLDROHFE 1077  
QY 787 FGR-----SKDYVVCRRNHACIGSVYLSLTOQALCGSYLFGDAPIRASYGP-----GN 834  
Db 1078 DGKYLFGSRGELAERSGH---IGLGIQSHQ---LGNLMIQQAIAKGNIGYVRFPSDGH 1132  
QY 835 QHMKTSTYFAEESD-----VRWDN-----NCLAGEITAGLPIVITPSKL 873  
Db 1133 EVHSPFONHSHSDSDEAGSPVDFSLYRIMWDGYEHPADYDGPQGGYPAPKGARDI 1192  
QY 874 YLNELRFPVQ-AEPSYADHESFTTEGQOARAFKSGHLLNLSVPVGVKFDRCSSHPNKYS 932  
Db 1193 YSYDIKGAQVIRNLNDNRS--TQORLADRFHNAAGSMLTQGVGDGPK-----RATRYS 1244  
QY 933 PMAAYICDAYRTISGTETTLTSHOETWTTDAFLHARHGVVVRGSMYASLTSNI-----EV 987  
Db 1245 PELDRSGNAEAFNGT-----ADIVKNIIGAAGEI 1274  
QY 988 YGHRGYEVRDASRGYGLSAGSRV 1010  
Db 1275 VGAG-----DAVQ--GISEGSNI 1290

RESULT 29

US-11-052-554A-171

; Sequence 171, Application US/11052554A  
; Publication No. US20050288866A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 171  
; LENGTH: 2340  
; TYPE: PRT  
; ORGANISM: Rickettsia prowazekii  
; US-11-052-554A-171

Query Match 3.7%; Score 193.5; DB 11; Length 2340;

Best Local Similarity 19.6%; Pred. No. 5.1e-05; Indels 405; Gaps 45;  
Matches 191; Conservative 109; Mismatches 270;

QY 55 SGTTFPS--AGBLTKNLDNSIAALPLSCFGLLGSFTVLGRHSLTFENIRTSTNGAAL 112  
Db 731 AGSTVLTDQTSBLTNN-DVVVNSNIITAGNNSKGLIFTGNG-GIT-GNI--GANGAAL 785  
QY 113 SNSAADGLFTIEG-FKELSFSCNCSLLAVLPRAATNKGSGTPTTTTPSPNGITYSKTDLL 171  
Db 786 QEVVFGTNTTIGGTANSQNFVAHSAANVVIITGLTTGALKYKDTGTIIAHGGVLGVDID-- 843  
QY 172 LLANKEFSYVNLVSDGGGAIDAKSLTVQGISKL CVFQENTAAQADGACQOVVTSFA-- 228  
Db 844 -FNKAGP-----ILGDGAMIDGSLVLCNGGVAGTLDP-----IGDGNVTQNICADNANSI 893  
QY 229 -----MANEAPIAPVANAVRGGGIAAQQDGGQVSSSTSTEDPVVVSFSR 274  
Db 894 STINIQDNTKNTVIAND--IFVDNIHFTNGGIL-----QLGGNLTTHN-----IDFGA 940  
QY 275 N-TAVFDFDGNVARGGGIYSYGNVAFNNKTLFLN-----NVA 312  
Db 941 NGGTLEFNGN-----NTYNLNAIIVNGQILNAFTNLKASDDTIGTVKIINIGQIG 992  
QY 313 SP-----VYIAAKQPTSGQASNTSNVYGD-----GGAIFC-- 342  
Db 993 TPQNFITQVNNKXLTIVSSVNSINFGDANSQILSAPVDQTIKFNNLNETGGIITLD 1052  
QY 343 -----KNGAQAGS--NNSGVSFDCG-----GVVFFSSNA 371  
Db 1053 SNGNLTISGNGIKLGSKNELSSLNIGKVTVTNDLDIQNTIHLNINNGALFDDQSLT 1112  
QY 372 AGK-----GGAIY----- 379  
Db 1113 SAKIKINIGTVAGGATYTLDAINDFNLTSGMVFQKDSILLELNKSSNTNDHTITLS 1172  
QY 380 -----AKLSVANGCPVQFLNAN-----DGAIVL----- 406  
Db 1173 ALDPGNQFGIILITDTNKLTIIDNNGNVAYTGLTANHMLKQLTTFASIDNGALKVGIN 1232  
QY 407 -----GESGELSADYDIIIDG 425  
Db 1233 VENTLNIKDIELNEVNAVLFNKNTYTTATGNGHVDQFGHAGVINLN--DDIIDIG 1289  
QY 426 NLKRTAKENAADVNGVTVSSQAISMGGKKIT-----TIRAKAGHQIL-----FNDPI 473  
Db 1290 SVTST-----GNVNG-----TLNFNGSGKVTVGLINNI VMLQAGAGDVLSASNGYSITE 1338  
QY 474 EMANGNNQ-----PAQSSKLLKINDGEGYTGDIVFANGSSTL----- 510  
Db 1339 IQGNGNNNTFAANSHTLTTDINKTGGQDLNLVFIINGGVSIGSISGANAAGVDIIINAGSVN 1398



Db	1629	LLSPRGLHDFWQKGYFNFLSGVYFVNNSSFSNATGSLNFVANKGIIFNGDNTID----	1684
Qy	80	SCFCNLLGSPFVLGRGHSLFENIRT--STNGAALNSAADGLFTTEGPKELSFNSCSNL	137
Db	1685	--FSKYOGALIFASNGVSNI--NITTLNATNGLSL--NAGLNNVSVQKGEICINLNC---	1736
Qy	138	LAVLPAATTNKGSGTPTT--TSTPSNGTIYSKTDLLLLNNEKPFYSNLSVGDGGAIDAKS	196
Db	1737	-----PTTKNSSPANSSVPTPNESLSVHANNFTFLGTIIS--NGAIDLSQ	1779
Qy	197	LTVOGISEKLCVFQEN--TAQADGGACQVVTFSAMANEAPAFVANVAGVGGGIAAVQDG	255
Db	1780	VTNNSVIGTTLNENATLOANN-----LTTTNAPNN-----	1810
Qy	256	QOGVSSSTSTEDPVVPSRNTAVFDCGNVARVGGIYSYGNVAPLNNKTKTLFNNVASPV	315
Db	1811	---ASNSTANIDGNFTLNOQATLSNAGSLNVGNFNSYGDLP-----NLSHSV	1857
Qy	316	YIAAQPTSGQASNTSNYDGGAI FCKNGAQGSNNSSGVSPFGEQGVFFSSNVAAGKG	375
Db	1858	-----SHAIINTQGT-----ATIMANNPLIQFN-----ASSKE	1886
Qy	376	GAIVAKKLSVANGCPVQFLNIANDGGAIVLGESEGLSLSADYGDIIIPDGNLAKRTA----	431
Db	1887	VGTY-----TLIDSAKAIYYGNNQIT-----GGSLDNLYLKLIALIDI	1925
Qy	432	--KENAADVNGVTYSSQAISSMGSGEKTTLIRAKAGHOI-----LFNDPIEMAN	477
Db	1926	NGKHWMTDNLGTNGQAVSVKQGLVVGFKDSQONQIYTSILYKNVKIAVNDPI----	1981
Qy	478	GNNOPAOSSKLLKINDGEGYTGDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQT	537
Db	1982	--NPOA-----PTLKQYIAQIQG-----VQSVDSIDQA	2008
Qy	538	GGG-----LYMEAGSTWDFVTPPPQPPAAPAQLITLSNLHL--S	575
Db	2009	CGNOAINLWKIPETKGSPLFAPYLESHSTKDLTT---TAGDIANTLEVIANPNPKND	2064
Qy	576	LSSLLANNAVTPPTNPPAQQSHPAVIGSTTAGSVTISGPIFFEDLDOTAYDRVDMGSLN	635
Db	2065	ATNIIQIINTYQO-----QMSRLAKLSDTSS--TFARSPFLERLEAKNKRADAIPN	2113
Qy	636	QKINVLKIQ-----LGTKPPANAPSD-----LTLGNEMPKYG	667
Db	2114	AMDVLKYQRNRVKNVWATVGGASPISGTGTLGINVGYDRFKGVGVG--YAAYG	2172
Qy	668	YQGSWKLAWDPNTANGPYTLKATWTKTYGNPGBPERVASLVPNSLWGSILDIRSAHSIQ	727
Db	2173	YSG-----PHANITQSGSSNVNVGYYSRAFTKRSELTWLSL--NETWGYNKTFINSYDPL-	2224
Qy	728	ASVDGRSYCRGLWVGVS-----NPFYHDDRAL-----GQGVRYI--SG-----	764
Db	2225	LSIINQSTRYDTWTTDAKINGYDFMFKDKSVIFKPQVGLSYYIIGLSGLGIMDDPIYN	2284
Qy	765	-----GYSLGANSYFGSSMFLAFTEVFGR	789
Db	2285	OPRANADPNKKSVMITNIALRSRHHYFNKNSYYFVIADV--GR	2324

RESULT 32

US-11-052-554A-179  
 ; Sequence 179, Application US/11052554A  
 ; Publication No. US2005028866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sachdeva, et al.  
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
 ; FILE REFERENCE: 30853/40359A  
 ; CURRENT APPLICATION NUMBER: US/11/052,554A  
 ; CURRENT FILING DATE: 2005-02-07  
 ; PRIOR APPLICATION NUMBER: US 60/589,227  
 ; PRIOR FILING DATE: 2004-07-20

```

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 179
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Shigella flexneri 2a str. 2457T
US-11-052-554A-179

Query Match          3.6%; Score 190; DB 11; Length 955;
Best Local Similarity 20.7%; Pred. No. 2.5e-05;
Matches 208; Conservative 127; Mismatches 348; Indels 320; Gaps 55;

Qy 30 IMPVQG-----IVDGETLTVSPFYVIGDSGTTVFSAGELTLKXLDNSAAUPL----- 79
Db 79 VMPAGNTPGNINLGNDVTVN-----VNDASG---YAKG-IIIGKKNSTANRLTVDDV 129
Qy 80 ---SCFG-NLLGSFTVLGRGHSHTFENIRSTNGAALSNSA--ADGLFTIEGPKELSFNS 133
Db 130 GQTSAGINILIGDYTHADLG---TGSTIKSNDGIIIGHSTLTATQFTIE-----N 178
Qy 134 CNSL-LAVLPAATY---NKGSTQPTTST-----PSNGTI-YSKTDLALLANNEKF 178
Db 179 SNGIGLTINDYGSVDLGGSKIKTDGSTGYVIGLGNNAANGAARFATDTLTI----- 232
Qy 179 SFYSNLVSGDGAIDAKSLTVQGISLKVQENTAQADGGACQVVTFSAMANEAPATFV 238
Db 233 -----DVQGSYAMGINVQKNSVVDLGTNSTIKTNGDNGHGLWFGQVSANALTVDV 283
Qy 239 ---ANVAVRGGIIAAVDGQ---QGVSSSTSTEDPVVSFSFSNTAVEFDGNVARVGGG 290
Db 284 TGAANGVEVRGGTTTGADSHISSAQGGGLVTSSDATINFSGAQR-----NS 334
Qy 291 IYSYGNVAFLNKGTFLFNNAVSPVYIAAQPTSGQASNTSNNYGDGGAIFCKNGAQAQS 350
Db 335 IFSGSGYCASQATAVIN-----MQNTDITVDRNGSLALGLWALSGG 377
Qy 351 NNSG-SVSGFDGEGVYVFFSSNVAAGKGAIIYA-----KGLSVANGCPVQFLRIAND 400
Db 378 RITGDSLAI TG-----AAGARG-IYAMTNSQIDLTSDLVDMSTPDMAIATQHD 426
Qy 401 GG--AIYLGESGELSLGADYGDIIIPDGNLKR-----TAKENADAVNG---VTYSS 445
Db 427 DGYAASRINASGRMLIN--GSVLSKGGLINLMDHPGSMVTGSSLSDNVNGKGLDVAMNN 483
Qy 446 QAI SNGSGGKLTTLR-----AKAGHQILFNDPIEMANGNNQPAQSSKLLKIND 493
Db 484 SVMNVTSNNSLDTLALSHSTVDPAHSGTAGTFTTLN--VENLSGNSTFFMRADVVGEGN 541
Qy 494 GEGVTGDIVFANGSSTLYQNVTIBQ-----GRVILR-----EKAKLSVNSLSQTGGS 540
Db 542 GVNREGDLLNTSGSAGNHVLAI RNQGESEATTGNEVLTVVKTTDGMAASFASQSVELGGY 601
Qy 541 LY--MEAGSTWDFV-----TQOP-----POQPPAANQLITLSLHLSLLSLLANNAV 585
Db 602 LYDVRKNGTNWELVYASGTVPEPTNPETPAQAQP-----FEDLDDTAYDRYD 638
Qy 586 TNP-PTNPAPQDSHPAVIGSTTA--GSVTISGPIF-----PEDLDDTAYDRYD 630
Db 639 VNPDPPTPEPATPKP-----TTTADAGGNLYLVNGYLLNVENRTLMQRMGDLRNQSKOGNI 694
Qy 631 WL-----GSKNINVLKQLQGTK-PAPANPSDLTLGN 661
Db 695 WLRSYGGSLDSFASGKLSGPDWGYSGI QFGDGKRLSDVMPLYVGLYIDSTHASPDYSGGD 754
Qy 662 EMPKYGYQGSWKLAWDNPNTANGPPT---LKAATWTKYNGPGERVASLVNSLWGSILD 718
Db 755 GTARSDYNGMYA---SYMAQNGFYSLDLVIKASRQKNSFH-----VLD 793
Qy 719 TRSAHSAIQASVDGRSYC-----RGLWVGVSNSFFYHDDRDLAQGYRYISGG 765
Db 794 SQNNGVNANGTANGMSISLEAGQRFNLSPGTGVGYIEPTQLTYSHQNMAM---KASNG 850

```



Qy	209	QENTAQADGGACQVVTSF-----SAMANEAPIAFVANVAGVGGIAAAVQDQGGVS	260
Db	421	LLTTAQ-DIGAVGYDSKFCHLLDAGKAMNGPASFPFGDFTADTKG-----	465
Qy	261	SSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF-----LNGKTLFLNNVASPV	315
Db	466	-----TSDIAYSF-RNDISGTGGLIKKGGSQLHGNNTYTKTIIIEGSSLVLYGNKNSDM	520
Qy	316	YIAAKQPT--SGQASNTSNNGDGAJFCNKGAAQAGSNNS-----GVSVPDGEGVWFSSN	369
Db	521	RVETKGALIYNGAAGSGSLN--SDGIVILADTQSGANETHIKGSLQLDCKGTLY----	574
Qy	370	VAACK-----GGAIYAKK-----LSVANCO-PVQFLERNIANDG	401
Db	575	TRLCKLLVDGTAIIGKLYMSARGKGAGYLNSTGRRVPFLSAAKIQDYSFPTTETDQ	634
Qy	402	GAIVLGSSEGLSADYGD---IIFGNLKRTRAK-----ENADVNGVTVSSQA	447
Db	635	GLLASLDSVEKTAGSE-GDTLSYVYRREGNAARTASAAHAHAPAGLKHAVQGGGSLNLELM	693
Qy	448	ISMG---SGGKITTLRAKAGHOILPNPDIEMANGNNQPAOSSKLLKLNIDEGYTGDIVFA	504
Db	694	VELDASESATPEVETAAA-----DRTDMPGIRPYGATFRAAAAQAHANAADGVRIFN	747
Qy	505	NGSSTLYQNVTIE---QGRIVLREKAKLSVNSISQTSGLYMEA-----GSTWD-----	550
Db	748	SLAATVYADSTAHAADMQGR---RLKAVSDGLDHNGTGLRVIAQTQDGGTWEQGGVE	802
Qy	551	-----FVTPQPPQPPAANQLITLSNHLJSLSLLANNAVTPPTNPAPQDSHPA	600
Db	803	GKMRGSTQVGIAAKTGENTTAA-----ATLGMGRSTWSNSA--NAKT-----DSISL	849
Qy	601	VIG-STTAGSVTISGPFIFFEDLDTAYDRYD-----WLGNO-----KINVLKQLGTKPP	650
Db	850	FAGIRHDAGDIIYKGLF-----SVGRYKNSISRSTGADERHAGSVNGTLMLQGLG	902
Qy	651	ANAP-----SDLTGLN-----EMPKYGYOGSWKLWDPNPTANNPGYTLKATWTKGYNP	699
Db	903	VNVPPAATGDLTVEGGLRYDLKQDAPEAKGS-ALGWSGNSLTEG-----	946
Qy	700	GPERVASIVPNSLMGSIID--IRSAHSAIQASVDGRSYCRGLWGSVGNPFYHDDRDLGQ	757
Db	947	---TLVGLAGLKLSQLPSDKAVLFATAGVERDLNGRDYT-----VTGFTGATAATGK	996
Qy	758	-----GYRISGYSIGANSYCGSSMFLAFTVEFGRSKDVVYCRSNHHACIGSVYLS	810
Db	997	TGARNMPTIRLVAG---LGADVFEFGWNGNLARYSYAG-SKQY-----GNHSGRVEGY--	1046
Qy	811	TOQALCGSYLFGDA-----FIRASYGFCNQ-----HMKTSYTFABESDVRWDN	853
Db	1047	--RFLGEGGGTGSATNDDDDVKKAATVAIAAYNNQOEINGPKAGETIYDIDEDCTIT-KK	1103
Qy	854	NCLAGEI-----GAGLPVITPSPKLYINELRPPVQAEFSYADHE	892
Db	1104	DATAADVBADDPKGLGLKKVVTNLTKTVNENKONVDAKVKAASE	1148



Matches 170; Conservative 110; Mismatches 293; Indels 219; Gaps 38;	
QY	23 GGGYAAEIMVPGIYDGETLTVSPYTVIGDPSGTVTSAGELTKLNLDNSTAALPLSCF 82
DB	284 GMYADSI-----TLIANEKGVGNAGTL-----EAAQLIVTSS 319
QY	83 GNLLGS----FTVLGRGHSHTPENIRTSNGAA---LNSA---ADGLFTTIEGPKELSPS 132
DB	320 GRIENSGRIATTADGTETASPTVLSIETTEKGAAGTFISNGGRIESKGLLVETGEDISLR 379
QY	133 NCNLSLAVLPAATTNKGSTPTTT--STPSNGTIYSKTDLLLNNKFPSPSYNLVSGDGA 191
DB	380 N-----GAVQNGSRPAITVLNAGHNLVIESKTN---VNNAKGS--ANLSA--GGR 424
QY	192 IDAKSLTVQIGSKLVQFQENTQAAGGACQVVTFSAMANEAPIAFVANAVAGVRGGIAA 251
DB	425 TTINDATIQAQSVSVSYSTKGDTELAGENTRIIAENVTLSNGS-----IGSAAVIEAKDTAH 480
QY	252 VODGQ-----QGVSSSTSTEDPVVSFSRNTAVEFDGNVAVRGVGGIYSYGNVAFLNNGKT 305
DB	481 IESGKPLSLETSTVASINIRLNNGNIKGGKQALLADDNITAKTNTLNTFENL--YVHTGKD 539
QY	306 LFLN--NVASPVVIAAKPTSGOASNTSNNGYDGGAI FCKNGAQAQ-----SNNSG 354
DB	540 LNLNVDKLSAASIHLKSDNAHICTSKLTASKDM-----GVEAGLLNVVTNLTNTSG 595
QY	355 SVSPDGEVYVFFSSNVAAKGG--AIYAKKLSVANCGPVQFLR--NTANDGAIYVIGESBEL 412
DB	596 NL-----HIQAAGNIQLRNTKLNAAKALETTALQCNIVSD--GLHAVSADGHV 642
QY	413 SLSADYGDIIIPDGNLKRTAKENADYNG-----VTVSSQAISM 450
DB	643 SLLAN--GNADFTGHNTLTAK---ADVNAGSVGKRLKADNTNITSSSGDITLVAGNGIQ 698
QY	451 GSG-----GKTTLRKAGHOILFNDPIEMANG-----NNQPAQSS----- 486
DB	699 GDGQKQNSNGKHISIKNGGNADLKNLVHAKSGALNIHSDRALSIENTKLESTHTHL 758
QY	487 ----KLLKINDGEGY-----TGDIVF-----ANG-----SSTLYQNVTI 516
DB	759 NAQHERVTLNQVDAYAHRHLSITGSQIWQNDKLP SANKLVANGVLALNARYSQIADNTTL 818
QY	517 EQGRIVLREKAKL-----SVN-----SLSQTCGSLYMEAGSTWDFVTPPPQQ 559
DB	819 RAGAINLTAGTALVKGKRNIMWSTVTKTLEDNAELKPLAGRLNIEAGS--GTLTIEP--- 873
QY	560 PPAANOLITLNLHLSLSLLANNAVTPNPTNPPAQDSH-----PAVIGST----- 605
DB	874 ---ANRISAHTDISITGGKLLLSAKGNGNAGAPSAQVSSLEAKGNIRLVGTETDLRSKI 930
QY	606 TAGS----VTISGPIPFEDLDDTDAYDRYWLGS---NQKINVLLKQLGTGKPPANAPSDL- 657
DB	931 TAGKNLVVATTKKLNIEAVNNSFSNVYFPTQKAELNOKSKELEQOIAQLKKSSPKSLKI 990
QY	658 -TLGNEPKYGY 668
DB	991 PTLQZERDRDLAF 1002
RESULT 38	
US-11-067-260-20	
; Sequence 20, Application US/11067260	
; Publication No. US20060051840A1	
; GENERAL INFORMATION:	
; APPLICANT: Arico, Maria	
; TITLE OF INVENTION: HETEROLOGOUS EXPRESSION OF NEISERIAL PROTEINS	
; FILE REFERENCE: 223002099501	
; CURRENT APPLICATION NUMBER: US/11/067,260	
; CURRENT FILING DATE: 2005-02-25	
; PRIOR APPLICATION NUMBER: 10/220,480	
; PRIOR FILING DATE: 2002-08-28	
; PRIOR APPLICATION NUMBER: PCT/IB01/00420	
; PRIOR FILING DATE: 2001-02-28	

Query Match	
3.6%; Score 188.5; DB 11; Length 1312;	
Beat Local Similarity 20.1%; Pred. No. Se-05;	
Matches 225; Conservative 140; Mismatches 421; Indels 331; Gaps 56;	
QY	60 PSAGSLTLKLNLDNSTAALPLSCFNGLLGCSFTVLGRGHSITF---ENIRTSNTGAAL---- 112
DB	268 YSGDKT---DEGIRLMQOQSDYGNL--SYHIRKNKMLFIFSTGNDAAQAPNTYALLPPY 321
QY	113 SNSAADGLFTIEG-----FKELSP-----SNCSLLAVLPAATTNKGSTPT 154
DB	322 EKDAQKGIITVAGVDRSGEKFKREMYGPGTEPLEYGSNHCGITAMWCLSDAPYEASVRPT 381
QY	155 TTS-----TPSNGTIYSKTDLLLNNKFPSPSYNLVSGDGAIDAKSLTVQIGSKLCVP 208
DB	382 RTNPQIAGTSFSAPIVTGTAAALL--QKYPMSN-----DNLRRT 420
QY	209 QENTQAAGGACQVVTSP-----SAMANEAPIAFVANAVAGVRGGIAA VODGQGVVS 260
DB	421 LUTTAQ--DIGAVGDSKFGWGLLDAGKAWNGPASPFPFGDTADTKG----- 465
QY	261 SSTSSTEDPVVSPSRNTAVEFDGNVAVRGVGGIYSYGNVAF-----LNNKTLFLNNVASPV 315
DB	466 ----TSDIAYSP-RNDISGTGLIKKGGSQLGHGNTVTGKTIIEGGSVLVYGNKNSDM 520
QY	316 YIAAKOPT--SQOASNTSNNGYDGGAI FCKNGAQAQSNNS-----GSVSFDGEGVFFSSN 369
DB	521 RVETKGLIYNGAASGGSLN--SDGIVVLADTDQSGANETVHIKGSLOLDGKGLY---- 574
QY	370 VAAGK-----GGAIYAKK-----LSVANGC-PVQFLRNANDG 401
DB	575 TRLGKLLKVDGTAAIIGKLYMSARGKAGYLNSTGRVFPFLSAAKIQDYSFTNLETDG 634
QY	402 GAIYLGESGELSLSADYGD---IIFDGNLKRITAK-----ENAAADVNGVTYSSQA 447
DB	635 GLLASLDSVEKTAGSE-GDTLSYVYRRGNAARTASAAAHSA PAGLKHAVEQCGSNLENLM 693
QY	448 ISMG---SGKITTIRAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFA 504
DB	694 VELDASESSATPETVETAAA-----DRDMPGIRPYGATPRAAAAVOHANAADGVRIFN 747
QY	505 NGSSSTLYQNVYIE---QGRIVLREKAKLSVNSLSQTCGSLYMEA-----GSTWD----- 550
DB	748 SLAATVYADSTAHAADMOGR-----RLKAVSDGLDHNGTGLRVIACTQDGGTWEQGVYE 802
QY	551 -----FVTPQPPQPPAANQLITLNLHLSLSLLANNAVTPNPTNPPAQDSHPA 600
DB	803 GMRGSGTQTVGIAAKTGENTTAA-----ATLCMGHSTWSENSA--NAKT-----DSISL 849
QY	601 VIG-STTAGSVTISGPIPFEDLDDTDAYDRYD-----WLGSNQ-----KINVLLKQLGTKPP 650
DB	850 FAGIRHDAGDIGYLKGLF-----SYGRYKNSISRSTGADEHAEGSVNGTLMQLGALGG 902
QY	651 ANAP-----SDLTIGN-----EMPKYGYQGSWKLANDPNTANNGPYTLKATWTKTYNP 699
DB	903 VNVPPAATGDLTVEGGLRYDLLKQDAPAEKGS--ALGWSGNSLSTEG----- 946
QY	700 GPERVASLPNSLWGSILD--IRSAHSAIQASVDGRSYCKRGLWVGSVNSFFYHDRDALGQ 757
DB	947 ----TLVGLAGLKLSQPLSDKAVLFATAGVERDLNGRDYT-----VTGFTCATATGK 996
QY	758 -----GYRIISGGYSLGANSYFGSSMFLAFTEVFGRSKDYVCRSNHHACIGSYLVS 810
DB	997 TGAENPHTELVLVAG---LGADVEFGWNGNGLARYSVAG--SKQY-----GNHSGRVG----- 1043



Qy 811 TQALCGSYLFGDAPIRASYGFGQNHMKTSYTPAESDVRDNNCLAGEIGAGLPVITP 870  
 Db 1044 -----VGYRFLGSGGG-----VAADIGAGLADALTA 1071  
 Qy 871 SKLYNELRPVQABFSAADHESFTEGDQA-RAPKSGHLNLSVPVGVKFDRCSSSTHPN 929  
 Db 1072 PLDHKDKGLQSLTLDSQVRKEKLAQAQGAETKYGNGDSLNTG---KLKDKVUS----- 1123  
 Qy 930 KYSFMAAYICDA-----YRTISGTEPTLLSHQETWTTDAPHLARHGYYVVR-----G 975  
 Db 1124 RPDFIRQIEVDGQLITLESQFQVVKQSHSALTAFQTEIQDSEHSGK--MVAKRQFRIG 1181  
 Qy 976 SMYASLTSNIEVYCHGRYEYDASRGYGLSAGSRVP 1012  
 Db 1182 DIAGEHTSFQKLPBGGRATYRGTAFGSD--DAGGKLTY 1217  
 RESULT 39  
 US-11-052-554A-6  
 ; Sequence 6, Application US/11052554A  
 ; Publication No. US2005028866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sachdeva, et al.  
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
 ; FILE REFERENCE: 30853/40359A  
 ; CURRENT FILING DATE: 2005-02-07  
 ; PRIOR FILING DATE: 2005-02-07  
 ; PRIOR FILING DATE: 2004-07-20  
 ; PRIOR FILING DATE: 2004-07-20  
 ; PRIOR FILING DATE: 2004-02-06  
 ; NUMBER OF SEQ ID NOS: 763  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 6  
 ; LENGTH: 949  
 ; TYPE: PRT  
 ; ORGANISM: Escherichia coli 0157:H7  
 US-11-052-554A-6

Query Match 3.6%; Score 187.5; DB 11; Length 949;  
 Best Local Similarity 20.7%; Pred. No. 3.7e-05;  
 Matches 203; Conservative 120; Mismatches 350; Indels 307; Gaps 53;  
 Qy 27 AAELMVPQG--IYDGETLTVSFPYTVIGDPSGTTVPFAGELTLKLNDSIAALPLSCFN 84  
 Db 52 AADKVVQAGETVNDG-TLTNHDNQIVFTGANGMTISTGLELGPDSSENTGGQWION--GG 108  
 Qy 85 LLGSFTVLGRHSUTPENIRTSNGAALSAAADGLFTIETGKLSFNSCNSLLAVLPAA 144  
 Db 109 IAGNTVTVTNGRQVLE-----GGTASDTVIRDG-----GGQSLNGLAVNTTL----- 151  
 Qy 145 TTNGSGSTPTTSTSPNGTIYSKTDLLLNNEKSPYSNLVSG-----DGGAI 192  
 Db 152 -NNKGBW-VHEGGVATGTIINRDG-----YQSVKSGGLATGTINTGAEQPD 198  
 Qy 193 DAKSLT---VQIGKLCVQENTAQADGACQVVTFSAMANEAPAFVANVAGVGGGI 249  
 Db 199 SDNSYTGKQVGTAEISTINKGRQ-----IILFSLGARDT-LIYAGDQSVHGRAL 249  
 Qy 250 -AAVDQDQGVSSSTSTEDPVVPSRNTAVFEDG-NVARVGGGIYSYGNVAFNLNGKTLF 307  
 Db 250 NTLNNGGYQYVHRD-----GLALNTVINEGGWQVVKAGG---AAGNTTINQNGE--- 295  
 Qy 308 LNNVAPVYIAAKPTSQASNTSNYGDGAIFCKNGAQA-GSNNSGSVDF-----GGGV 363  
 Db 296 -----LRVHAGGEATVATQN---TGGALVTSTAAVTGIVNRLNGFNFTVNGKADGV 342  
 Qy 364 VFFSNNVAAGKGGAIYAKLSVANCGPVQVFLRNIAINDGGAIIYLGESGEL-----SL 414  
 Db 343 VLES-----GG-----RLDVLSEHSAQ--NTLVDDGGTLAVSAGGKGVTSVITTSAGAL 388

Qy 415 SADYGDIIIFDGNLKRITAKENAADVNGVTVSQAISMGSCKIT-----T 458  
 Db 389 IADSG-----ATVEGTNASOKPSIDGTSQOASGLLLENGGSFTVNAGQAGNTTVGHRGT 443  
 Qy 459 LRAKAGHQIILFNDPIEMANGNQPASQSKLLKNDGEGYTGDIIVFANGSSSTLYQNVTIEQ 518  
 Db 444 LTLAAGGSL-----SRTQLSKGASMV-LNGDVVSTGDIV--NAGEIRFPDNTQTPN 491  
 Qy 519 ---GRIVLREKA-----KLSVNSLSOTGGSLYMEA---GSTWDFVTPVPPQPPPAANQLI 567  
 Db 492 AALSRAVAKNSPVTPHKLTTLNLTGGGTINMRVLDSN-----ASDQLV 538  
 Qy 568 TLSNLHLSSLANNAVTPNPPPAQDHPAIVGSTTAGSVTISGPIPFEDLDDTAYD 627  
 Db 539 INGGQATGKTWL-----APT-----VGNSTLGVAT----- 564  
 Qy 628 RYDWLGSNQKINVLKQLGTPKANAPASDLTLGNEMPKYGVQGSKWLAWDPNTANNPYT 687  
 Db 565 -----TGQIRVVDQAQNGATTE-----EGAPALS-RPLQAGAFNYT 599  
 Qy 688 LKATWTKTGYNPGPVRVASLVPNSLMGSIID-----IRSAHSAIQASVDGRSYCRGLW 740  
 Db 600 LNRDSDDEWYLRGENAVRAEVP--LYTSMLTQAMDYDRILLAGSRSHQTVGNNGENSVRLS 657  
 Qy 741 VSGVSNFPYHDDRALGCGYRIYISGGYSLGANSYFGSSMFG-----LATFEVGRSKDY 793  
 Db 658 IQG--GHLGHDN-----NGGIARGATPE--SSGSYGVRLLEGDLRTEVAGMSLT 704  
 Qy 794 VVCRSHHACI-----GSVYLSTQALCGSYLFGDAFTRASYGP-----GNQH 836  
 Db 705 GVTGAAGSHSVVDKDDGSRAGTV--RDDAGSLGGYL--NLVTSGLWADIVAQOTRH 759  
 Qy 837 -MKTSTYTPAESDVRDNNCLAGEIGAGLPVITPISKLYNELRPVQABFSAADHESFT 895  
 Db 760 SMKAS---SDNDFRAGWGMWLSLETGLPFSITDNL-----LEPQLQ-----YT 802  
 Qy 896 BEG-----DQAPAFKSH 908  
 Db 803 WQGLSLDDGQDNAGYVKFGH 822  
 RESULT 40  
 US-11-052-554A-284  
 ; Sequence 284, Application US/11052554A  
 ; Publication No. US2005028866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sachdeva, et al.  
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
 ; FILE REFERENCE: 30853/40359A  
 ; CURRENT FILING DATE: 2005-02-07  
 ; PRIOR FILING DATE: 2005-02-07  
 ; PRIOR FILING DATE: 2004-07-20  
 ; PRIOR FILING DATE: 2004-07-20  
 ; PRIOR FILING DATE: 2004-02-06  
 ; NUMBER OF SEQ ID NOS: 763  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 284  
 ; LENGTH: 1343  
 ; TYPE: PRT  
 ; ORGANISM: Escherichia coli 0157:H7  
 US-11-052-554A-284

Query Match 3.5%; Score 186.5; DB 11; Length 1343;  
 Best Local Similarity 19.5%; Pred. No. 7.2e-05;  
 Matches 174; Conservative 123; Mismatches 310; Indels 285; Gaps 43;  
 Qy 15 AYSCCSLNGGGYAAEIMVPGQIYDGETLTVSFPYTVIGDPSGTTVPFAGELTLKLNDSI 74  
 Db 494 SPDTVGIVNVPFGSGIV---NVNGATLNTSTGYGFIGNAGSKGIVNISTDSLWNLKTS 550  
 Qy 75 AALPLSCFGLNLSGFTVLRGRGH-SLTIPENIRTSNGAALSAAADGLFTIETGKLSFNS 133

Db 551 TNAQLQVG-----VLGTGELNITTCGIVVARDTQIALNDKSKGDRVVDG----- 595

QY 134 CNSLLAV--LPAATTNKGSTPTTTST-----PSNGTI-----YSKTDL 170

Db 596 QNSLLETNNYVGTSTGTLTLTNSGTLNVEGGEVVLGVFEPVAVGTLNIGAAGHGAADA 655

QY 171 -LLLNNKPSFYSN-----LVSGD-----GGAI 192

Db 656 GPITNATKVFEGSGEGVFVFNHTNNSDAGYQVDMLTIGDDKDKVHDAHTVFNAGNTY 715

QY 193 DAKSLTVQGISKLCVFOENTAOAG---GACQVTFPSAMANEAPIAFVANVAG----- 243

Db 716 SKTLVNDGLLIA-----SHTADGVTGMSSEVTI---ASPCTLDLASTNSAGDYTLT 767

QY 244 --VRGGIAAVQDQGGVSSSTSTEDFVVSFERNTAVERFDG----- 282

Db 768 NALKGDGLMRVQ-----LSSDKMFGFTHATGTEPAGVAQLKMDSTFLERDNTAAL 818

QY 283 -----NVARVGGGIYSGNVAFNNGKTLFLNNVASPV-----YIAAKOPTSGQAS 328

Db 819 THAMLOSDIENTTSNVNQGSIQGLA--MNGGTLIFDTIPAAATLARGYISVDTLVVVGASD 877

QY 329 NT--SNNY--GDGGAIF--CKNGAQAGSNN--GSVSFDGE--GVVPFSSNVAAGKG 375

Db 878 YTWKGRNYQVNGRQDVLIGVPKPNWDPMANPLTTLNLEHDDNHVGVQLVRAQTVIGSG 937

QY 376 GAIYAKKLSVANGCPVQFLNIAND---GGAIYLGESGELSLSADYD-----DII 422

Db 938 GSL-----TLRDQGEVEADKTLHIAQNGTVVAEGDYGFRLLTTPAGDGLY 983

QY 423 FDGNLKRKAKENADVNTVSSQAI SMGSGKITTILRAKAGHQ--ILFND--PIEMANG 478

Db 984 VNYGLK-----ALNTHGGQKLTAEHGGAYGATADMSAKIGGEDLAINTVRQVSLSG 1037

QY 479 NNQPAQSSKLLKI-----NDGEGVTGD-----IVFANGSSTLYQNVITIEQGRIVLR 524

Db 1038 QND-YQGATVQMGTLRTDADGALGNTRLENI SNAALVDLNGSTQVTFETFGQGSTVLF 1096

QY 525 EKAKLSVN-----SLSQTGGSLYMEA-GSTWDFVTFPPQPPQPPAANQITLS 570

Db 1097 KEGSLTVNKGGIQSGELTGGGNLNVGTGLAVEGLNARYNALT VSPNAEVS LDTQGLG 1156

QY 571 NLHLSLSSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFELDDTAYDRYD 630

Db 1157 RGNIANDGLLTLKNVTGELRN---SISGKGIVSATARTDVELDG-----DNSR 1201

QY 631 WLGSNQKINVLQLGKTPPANAPSDLTLGNEMPKYGYQGSWKLAWDPTANNKGPVTL-- 688

Db 1202 FVG---QFNI-----DTGSALSV-NEQKNLG-----DASVNNGLLITIST 1237

QY 689 KATWTKTYNPGPVERVASLVPSNLWGS-----ILDIRSAHSAIQASVD 731

Db 1238 ERSWAMT-----HSISGSGDLTKLGTGILTLNNDSSAYQGTTD 1275

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:11:34 ; Search time 190 Seconds  
(without alignments)  
2340.270 Million cell updates/sec

Title: US-10-701-844-2  
Perfect score: 5267  
Sequence: 1 MQTSPHKFLLSILAYSCS.....VEYRDASRGYGLSAGSRVRF 1012

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 21.\*

1: Geneseqp1380s.\*  
2: Geneseqp1390s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	100.0	1012	2 AAY16735	Aay16735 C. tracho
2	5131.5	97.4	1013	5 ABG91021	Abg91021 Chlamydia
3	5131.5	97.4	1013	7 ADD43798	Add43798 Chlamydia
4	5131.5	97.4	1013	7 ADD43722	Add43722 Chlamydia
5	5131.5	97.4	1013	7 ADD43718	Add43718 Chlamydia
6	5131.5	97.4	1013	7 ADD43708	Add43708 Chlamydia
7	5131.5	97.4	1013	7 ADD43714	Add43714 Chlamydia
8	5131.5	97.4	1013	7 ADD43786	Add43786 Chlamydia
9	5131.5	97.4	1013	7 ADD43788	Add43788 Chlamydia
10	5131.5	97.4	1013	7 ADD43710	Add43710 Chlamydia
11	5131.5	97.4	1013	7 ADD43706	Add43706 Chlamydia
12	5131.5	97.4	1013	9 ADW29027	Adw29027 C. trachom
13	5131.5	97.4	1013	9 AEA19078	Aea19078 Chlamydia
14	5131.5	97.4	1013	9 AEA19002	Aea19002 Chlamydia
15	5125.5	97.3	1013	9 AEA19080	Aea19080 Chlamydia
16	5123.5	97.3	1013	9 AEA19006	Aea19006 Chlamydia
17	5123.5	97.3	1013	9 AEA19090	Aea19090 Chlamydia
18	5123.5	97.3	1013	9 AEA19010	Aea19010 Chlamydia
19	5121.5	97.2	1013	9 AEA19000	Aea19000 Chlamydia
20	5121.5	97.2	1013	9 AEA18998	Aea18998 Chlamydia
21	5112.5	97.1	1013	9 AEA19014	Aea19014 Chlamydia
22	5106.5	97.0	1013	2 AAY16737	Aay16737 C. tracho
23	5106.5	97.0	1013	2 AAY16738	Aay16738 C. tracho
24	5090	96.6	1006	4 ABG83207	Abg83207 Protein e

## ALIGNMENTS

### RESULT 1

AAY16735

ID AAY16735 standard; protein; 1012 AA.

XX AC AAY16735;

XX DT 21-JUL-1999 (first entry)

XX DE C. trachomatis LGV L2 HMW protein.

XX KW Chlamydia; high molecular weight protein; HMW protein; urethritis;

KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;

KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;

KW salpingitis; tubal occlusion; infertility; cervical cancer;

KW arteriosclerosis; atherosclerosis.

XX OS Chlamydia trachomatis.

XX PN W09917741-A1.

XX PD 15-APR-1999.

XX PF 01-OCT-1998; 98WO-US020737.

XX PR 02-OCT-1997; 97US-00942596.

XX PA (ANTE-) ANTEX BIOLOGICS INC.

XX PI Jackson JW, Pace JL;

XX DR WPI; 1999-287659/24.

XX DR N-PSDB; AAX60539.

XX PT New Chlamydia protein useful for treating conjunctivitis, urethritis and cervical cancer.

XX PS Claim 4; Fig 3; 141pp; English.

XX CC The invention relates to an isolated Chlamydia species high molecular weight (HMW) protein having an apparent mol. wt. of 105-115 kD as determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can be used for preventing, treating or ameliorating a disorder related to Chlamydia e.g. bacterial infection, conjunctivitis, urethritis, lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis, CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical cancer, infertility, arteriosclerosis and atherosclerosis. The products CC can also be used for detection and diagnosis. The present sequence

Abb94178 Chlamydia  
Aab13633 C. tracho  
Aag83201 Protein e  
Abb94172 Chlamydia  
Aab13639 C. tracho  
Add42756 Chlamydia  
Aay37238 Chlamydia  
Aay16739 C. tracho  
Aay16751 Chlamydia  
Aay16752 Chlamydia  
Abb90527 Chlamydia  
Aay96274 Chlamydia  
Aay69369 Amino aci  
Aaw84222 Chlamydia  
Aay94327 Chlamydia  
Abp56019 Chlamydia  
Abb98228 Chlamydia  
Abu66284 C. psitta  
Aaw88421 Chlamydia  
Aaw88418 Chlamydia  
Abb90573 Chlamydia

25 5090 96.6 1006 5 ABB94178  
26 5084 96.5 982 3 AAB13633  
27 5084 96.5 982 4 AAG83201  
28 5084 96.5 982 5 ABB94172  
29 5083 96.5 1006 3 AAB13639  
30 3336.5 63.3 670 7 ADD42756  
31 2593 49.2 524 2 AAY37238  
32 2547 48.4 505 2 AAY16739  
33 2350 44.6 458 2 AAY16751  
34 1735 32.9 325 2 AAY16752  
35 1365.5 25.9 973 5 ABB90527  
36 1359.5 25.8 973 3 AAY96274  
37 1165 22.1 918 3 AAY69369  
38 1143 21.7 918 2 AAW84222  
39 1128 21.4 928 3 AAY94327  
40 1122 21.3 926 5 ABP56019  
41 1122 21.3 926 5 ABB98228  
42 1122 21.3 926 6 ABU66284  
43 1121 21.3 928 2 AAW88421  
44 1101 20.9 928 2 AAW88418  
45 1091 20.7 928 5 ABB90573

CC represents a C. trachomatis HMW protein  
XX Sequence 1012 AA;

```
Query Match      100.0%; Score 5267; DB 2; Length 1012;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQTSFHKFLSMILAYSCSLNGGGYAAIRIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP 60
Db 1 MQTSFHKFLSMILAYSCSLNGGGYAAIRIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP 60

Qy 61 SAGELTKNLNDNSIAALPLSCFNLGSLTFVLRGHSITFENIRTSNGAALSNSAADGL 120
Db 61 SAGELTKNLNDNSIAALPLSCFNLGSLTFVLRGHSITFENIRTSNGAALSNSAADGL 120

Qy 121 FTIEGFKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180
Db 121 FTIEGFKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180

Qy 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTQAADGGACQVTSFSAMANEAPAFVAN 240
Db 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTQAADGGACQVTSFSAMANEAPAFVAN 240

Qy 241 VAGVRGGGIAAVDQGGQGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
Db 241 VAGVRGGGIAAVDQGGQGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300

Qy 301 NNGKTLFLNNVSPVYIAAKPTSQASNTSNNGDGAIFCKNGAQAGSNNSGSVSPDG 360
Db 301 NNGKTLFLNNVSPVYIAAKPTSQASNTSNNGDGAIFCKNGAQAGSNNSGSVSPDG 360

Qy 361 EGVVFFSNVAAAGKGAAYAKKLAVANGCPVQFLNIANDGGAAYLGSBELSLADYGD 420
Db 361 EGVVFFSNVAAAGKGAAYAKKLAVANGCPVQFLNIANDGGAAYLGSBELSLADYGD 420

Qy 421 IIFDGNLKTAKENADVNGTVSSQALSMGSGKITTILRAKAGHOILFNDPIEMANGW 480
Db 421 IIFDGNLKTAKENADVNGTVSSQALSMGSGKITTILRAKAGHOILFNDPIEMANGW 480

Qy 481 QPAOSSKLLKNDGEGYTDIVPANGSTLYQNTVIEQRIVLREKAKLSVNSLSOTGGS 540
Db 481 QPAOSSKLLKNDGEGYTDIVPANGSTLYQNTVIEQRIVLREKAKLSVNSLSOTGGS 540

Qy 541 LYMEAGSTWDFVTPQPPQPPAAANQLITLSNLHLSLSLLANNAVTPNTPPAQDSHPA 600
Db 541 LYMEAGSTWDFVTPQPPQPPAAANQLITLSNLHLSLSLLANNAVTPNTPPAQDSHPA 600

Qy 601 VIGSTAGSVTISGPIPEDDDDTAYDRYDNLGSKINVLKQLGTPKPPANAPSDLTIG 660
Db 601 VIGSTAGSVTISGPIPEDDDDTAYDRYDNLGSKINVLKQLGTPKPPANAPSDLTIG 660

Qy 661 NEMPKYGYQGSWKLAWDPNTANNGPYTLKATWTKTGNYPGPVERVASLVPNSLWGSILDIR 720
Db 661 NEMPKYGYQGSWKLAWDPNTANNGPYTLKATWTKTGNYPGPVERVASLVPNSLWGSILDIR 720

Qy 721 SAHSAIQASVDGRSVCRLWTVGSVSNFFYHDDRDALGQGYRIVISGYSLGANSYFGSSMFG 780
Db 721 SAHSAIQASVDGRSVCRLWTVGSVSNFFYHDDRDALGQGYRIVISGYSLGANSYFGSSMFG 780

Qy 781 LAFTEVFRSDYVYVCRSNHHACISVYLSLSTQALCGSYLPGDAFIRASYFGQHMKTS 840
Db 781 LAFTEVFRSDYVYVCRSNHHACISVYLSLSTQALCGSYLPGDAFIRASYFGQHMKTS 840

Qy 841 YTPAESDVRWNNCLAGEIGAGLPITVTPSKLYLNELRPPVQAEFSVADHESFTEGDQ 900
Db 841 YTPAESDVRWNNCLAGEIGAGLPITVTPSKLYLNELRPPVQAEFSVADHESFTEGDQ 900

Qy 901 ARAFKSGHLLNLSVPVGVKFCRCSSTHPNKYSFMAAYICDARYTISGTETLLSHQETWT 960
Db 901 ARAFKSGHLLNLSVPVGVKFCRCSSTHPNKYSFMAAYICDARYTISGTETLLSHQETWT 960

Qy 961 TDAFHARHGTVVRGSMYASLTSNIEVYGHGRYEDASRGYGLSAGSRVRF 1012
```

```
Db 961 TDAFHARHGTVVRGSMYASLTSNIEVYGHGRYEDASRGYGLSAGSRVRF 1012

RESULT 2
ABG91021
ID ABG91021 standard; protein; 1013 AA.
XX AC ABG91021;
XX 29-NOV-2002 (first entry)
XX Chlamydia trachomatis outer membrane protein G protein.
XX Gram-negative bacterial bleb; PorB; Outer membrane protein;
XX Chlamydia trachomatis infection; Chlamydia pneumoniae infection;
XX protective antigen; antibacterial; vaccine; gene; ds.
XX Chlamydia trachomatis.
XX WO200262380-A2.
XX 15-AUG-2002.
XX 08-FEB-2002; 2002WO-EP001356.
XX 08-FEB-2001; 2001GB-00003169.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Berthet FJ, Lobet Y, Poolman J, Verlant VGCL;
XX WPI; 2002-657510/70.
XX N-PSDB; ABS67342.
XX Novel gram-negative bacterial bleb presenting on its surface PorB outer
XX membrane protein from Chlamydia trachomatis or protective antigen from
XX Chlamydia pneumoniae, useful for preventing Chlamydia infection.
XX Disclosure; Page 10; 75pp; English.
XX The present invention relates to a new gram-negative bacterial bleb
XX presenting on its surface the PorB outer membrane protein from Chlamydia
XX trachomatis, or a protective antigen from C. pneumoniae. The invention is
XX useful for preventing C. trachomatis or C. pneumoniae infection in a
XX host. The present nucleic acid sequence represents a Chlamydia
XX trachomatis gene of the invention
XX Sequence 1013 AA;

Query Match      97.4%; Score 5131.5; DB 5; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

Qy 1 MQTSFHKFLSMILAYSCSLNGGGYAAIRIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP 60
Db 1 MQTSFHKFLSMILAYSCSLNGGGYAAIRIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP 60

Qy 61 SAGELTKNLNDNSIAALPLSCFNLGSLTFVLRGHSITFENIRTSNGAALSNSAADGL 120
Db 61 SAGELTKNLNDNSIAALPLSCFNLGSLTFVLRGHSITFENIRTSNGAALSNSAADGL 120

Qy 121 FTIEGFKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180
Db 121 FTIEGFKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180

Qy 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTQAADGGACQVTSFSAMANEAPAFVAN 240
Db 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTQAADGGACQVTSFSAMANEAPAFVAN 240

Qy 241 VAGVRGGGIAAVDQGGQGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
Db 241 VAGVRGGGIAAVDQGGQGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
```

QY 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYYGDDGGAIFCKNGAQ-AGSNNSGSVSPD 359  
DB 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDDGGAIFCKNGAQAGSNNSGSVSPD 360  
QY 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLRNANDGGAIFYLGESGELSADYD 419  
DB 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLGNANDGGAIFYLGESGELSADYD 420  
QY 420 DIIFDGNLKRKTAKENAADVNGVTSSQAISMGGGKITTLLRAKAGHQLFNDPIEMANGN 479  
DB 421 DIIFDGNLKRKTAKENAADVNGVTSSQAISMGGGKITTLLRAKAGHQLFNDPIEMANGN 480  
QY 480 NOPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQTCG 539  
DB 481 NOPAQSSSEPLKINDGEGYTGDIIVFANGNSTLYQNTYIEQGRIVLREKAKLSVNSLSQTCG 540  
QY 540 SLYMEAGSTWDFVTPOPPQPPAANQLITLSNLHLSSLLANNVTPPTPPAQDSHP 599  
DB 541 SLYMEAGSTLDFVTPOPPQPPAANQLITLSNLHLSSLLANNVTPPTPPAQDSHP 600  
QY 600 AVIGSTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKINVLKQLGTGTPPANASDLTL 659  
DB 601 AIIGSTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKIDVLKQLGTGTPPANASDLTL 660  
QY 660 GNEMPKYGYQGSWKLAWDPNTANGPYTLKATWTKTYNPGPERVASLVPNSLWGSILDI 719  
DB 661 GNEMPKYGYQGSWKLAWDPNTANGPYTLKATWTKTYNPGPERVASLVPNSLWGSILDI 720  
QY 720 RASHAIOASVGRSVCRLGVSGVSNPFPHDRDALGQGVRYISGGYSLGANSYFGSSMF 779  
DB 721 RASHAIOASVGRSVCRLGVSGVSNPFPHDRDALGQGVRYISGGYSLGANSYFGSSMF 780  
QY 780 GLAFTVFGRSKDYVVVCRSNHACIGSVYLSLSTQALCGSYLFGDAFIRASYGFGNQHMKT 839  
DB 781 GLAFTVFGRSKDYVVVCRSNHACIGSVYLSLSTQALCGSYLFGDAFIRASYGFGNQHMKT 840  
QY 840 SYTFABESVRMNNCLAGIAGLPIVITPSKLYLNELRPPVQAFPSYADHESFTEGD 899  
DB 841 SYTFABESVRMNNCLAGIAGLPIVITPSKLYLNELRPPVQAFPSYADHESFTEGD 900  
QY 900 QARAPKSGHLNLSVPVGVKFRDCSSTHPNKYFMAAYICDAVRTISGTTLLSHQETW 959  
DB 901 QARAPKSGHLNLSVPVGVKFRDCSSTHPNKYFMAAYICDAVRTISGTTLLSHQETW 960  
QY 960 TTDAPFLARHGVVVRGSMVASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF 1012  
DB 961 TTDAPFLARHGVVVRGSMVASLTSNIEVYGHGREYRDTSRGYGLSAGSRVRF 1013

RESULT 3

ID ADD43798 standard; protein; 1013 AA.

AC ADD43798;

15-JAN-2004 (first entry)

Chlamydia trachomatis immunogenic protein, SEQ ID NO 93.

immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
gene therapy; antibacterial.

Chlamydia trachomatis.

W02003049762-A2.

19-JUN-2003.

12-DEC-2002; 2002WO-IB005761.

12-DEC-2001; 2001GB-00029732.

06-AUG-2002; 2002GB-00018233.

14-AUG-2002; 2002GB-00018924.

(CHIR-) CHIRON SPA.

Grandi G, Ratti G;

WPI; 2003-532882/50.

N-PSDB; ADD43799.

New immunogenic composition having a protein or encoding nucleic acid,  
useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
infection.

Claim 6; SEQ ID NO 93; 164pp; English.

The invention relates to a novel immunogenic composition comprising a  
protein or nucleic acid, and an adjuvant, where the protein or nucleic  
acid comprises any of 131 fully defined amino acid or nucleotide  
sequences given in the specification, or has 50% or greater sequence  
identity to it, or their fragments. The protein and/or nucleic acid of  
the immunogenic composition is useful in the manufacture of a medicament  
for the treatment or prevention of infection due to Chlamydia  
trachomatis. The infection is treated or prevented by the medicament  
eliciting an immune response which is specific to a C. trachomatis  
elementary body, or for neutralising C. trachomatis elementary bodies,  
hence the immunogenic composition can be used in creating a vaccine. The  
immunogenic compositions can also be used for the diagnosis of C.  
trachomatis infection. The nucleic acids of the immunogenic compositions  
can be used to treat disorders by gene therapy. The immunogenic  
compositions have antibacterial activity. This sequence represents one of  
the 131 C. trachomatis proteins with immunogenic properties of the  
invention.

Sequence 1013 AA;

Query Match 97.4%; Score 5131.5; DB 7; Length 1013;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 MOTSFHKPLSLMILAYSCCLNGGGVYAAIMVPOQGIYDGETLTVSPYTVIGDPSGTTVF 60

DB 1 MOTSFHKPLSLMILAYSCCLNGGGVYAAIMVPOQGIYDGETLTVSPYTVIGDPSGTTVF 60

QY 61 SAGELTLKLNLSIAALPLSCFNLIGSFTVLGRGHSLTFFENIRTSNGAALSNSAADGL 120

DB 61 SAGELTLKLNLSIAALPLSCFNLIGSFTVLGRGHSLTFFENIRTSNGAALSNSAADGL 120

QY 121 FTIEGPKLSFSCNCSLLAVLPAATNKGSOPTTTSTPSNGTIYKTDLLLNNEKPSF 180

DB 121 FTIEGPKLSFSCNCSLLAVLPAATNKGSOPTTTSTPSNGTIYKTDLLLNNEKPSF 180

QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTFSAMANEAPIAFVAN 240

DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTFSAMANEAPIAFVAN 240

QY 241 VAGVRGGIAAVODGQGGVSSSTSTEDPVVPSFRNTAVEFDGNVAVRGGGIYSGNVAF 300

DB 241 VAGVRGGIAAVODGQGGVSSSTSTEDPVVPSFRNTAVEFDGNVAVRGGGIYSGNVAF 300

QY 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYYGDDGGAIFCKNGAQ-AGSNNSGSVSPD 359

DB 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDDGGAIFCKNGAQAGSNNSGSVSPD 360

QY 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLRNANDGGAIFYLGESGELSADYD 419

DB 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLGNANDGGAIFYLGESGELSADYD 420

QY 420 DIIFDGNLKRKTAKENAADVNGVTSSQAISMGGGKITTLLRAKAGHQLFNDPIEMANGN 479

DB 421 DIIFDGNLKRKTAKENAADVNGVTSSQAISMGGGKITTLLRAKAGHQLFNDPIEMANGN 480

QY 480 NOPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQTCG 539

481	DB	NQPAQSGBEPLKINDGEGVTGDIVFANGNSTLYQNVTTIISQGRIVLREKAKLSVNSLSQGTG	540
540	QY	SLYMEAGSTWDFVTVPQPQPQPPAAANQLITLSNLHLSSLSLLANNAVTNPPTNPPAQDSDHP	599
541	DB	SLYMEAGSTLDFVTVPQPQPQPPAAANQLITLSNLHLSSLSLLANNAVTNPPTNPPAQDSDHP	600
600	QY	AVIGSTTAGSVTTISGPIIFEPEDDLDDTAYDRYDWLGSNQKINTVLKLOLGTKPPANARSDLTLL	659
601	DB	AIIGSTTAGSVTTISGPIIFEPEDDLDDTAYDRYDWLGSNQKIDVLKLOLGTQPSANARSDLTLL	660
660	QY	GNEMPKYGYQGSWKLANDPNTANNPGYTLKATWTXTGYNPGPERSVASLVPNSLWGSILDI	719
661	DB	GNEMPKYGYQGSWKLANDPNTANNPGYTLKATWTXTGYNPGPERSVASLVPNSLWGSILDI	720
720	QY	RSAHSATQASVDGRSYRCGLWVSGVSNFFYHDDRDLGQGYRIISGYSILGANSYFGSSMF	779
721	DB	RSAHSATQASVDGRSYRCGLWVSGVSNFFYHDDRDLGQGYRIISGYSILGANSYFGSSMF	780
780	QY	GLATFTEVFGRSKDYVVCRSNHHACITGSVYLSLTOQALCGSYLFGDAFIRASVYFGNQHMKT	839
781	DB	GLATFTEVFGRSKDYVVCRSNHHACITGSVYLSLTKQALCGSYLFGDAFIRASVYFGNQHMKT	840
840	QY	SYTFAESDVRWDDNNCLAGETGAGLPDIVITP9SKLYLNELRPPVQAPRPSYADHESFTEBGD	899
841	DB	SYTFAESDVRWDDNNCLVGEIGVGLPIVITP9SKLYLNELRPPVQAFESYADHESFTEBGD	900
900	QY	QARAPKSHGLIANLSPVPGVKFDRCSSTHPNKYSFMAAVICDAYRTISGTHTTLLSHQRTW	959
901	DB	QARAPRSGHLMNLSPVPGVKFDRCSSTHPNKYSFMAAVICDAYRISGTHTTLLSHQRTW	960
960	QY	TTDAFHILARHGVVVRGSGMYASLT9NIEVYGHGRYEYRDASRGYGLSAGSRVRF	1012
961	DB	TTDAFHILARHGVIVRGSGMYASLT9NIEVYGHGRYEYRDTSRGYGLSAGSKVRF	1013

## RESULT 4

ADD43722	
ID	ADD43722 standard; protein; 1013 AA.
XX	
XX	ADD43722;
XX	
DT	15-JAN-2004 (first entry)
XX	
DE	Chlamydia trachomatis immunogenic protein, SEQ ID No 17.
XX	
KW	immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW	gene therapy; antibacterial.
XX	
OS	Chlamydia trachomatis.
XX	
PN	WO2003049762-A2.
XX	
PD	19-JUN-2003.
XX	
PP	12-DEC-2002; 2002WO-IB005761.
XX	
PR	12-DEC-2001; 2001GB-00029732.
PR	06-AUG-2002; 2002GB-00018233.
PR	14-AUG-2002; 2002GB-00018924.
XX	
PA	(CHIR-) CHIRON SPA.
XX	
PI	Grandi G, Ratti G;
XX	
DR	WPI; 2003-532882/50.
DR	N-PSDB; ADD43723.
XX	
PT	New immunogenic composition having a protein or encoding nucleic acid,
PT	useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT	infection.
XX	
PS	Claim 6; SEQ ID NO 17; 164pp; English.
XX	

The invention relates to a novel immunogenic composition comprising a protein or nucleic acid, and an adjuvant, where the protein or nucleic acid comprises any of 131 fully defined amino acid or nucleoside sequences given in the specification, or has 50% or greater sequence identity to it, or their fragments. The protein and/or nucleic acid of the immunogenic composition is useful in the manufacture of a medicament for the treatment or prevention of infection due to Chlamydia trachomatis. The infection is treated or prevented by the medicament eliciting an immune response which is specific to a C. trachomatis elementary body, or for neutralising C. trachomatis elementary bodies, hence the immunogenic composition can be used in creating a vaccine. The immunogenic composition can also be used for the diagnosis of C. trachomatis infection. The nucleic acids of the immunogenic compositions can be used to treat disorders by gene therapy. The immunogenic compositions have antibacterial activity. This sequence represents one of the 131 C. trachomatis proteins with immunogenic properties of the invention.

Sequence 1013 AA;

Query Match	97.4%	Score 5131.5	DB 7	Length 1013
Best Local Similarity	97.0%	Pred. No. 0		
Matches 983	Conservative 18	Mismatches 11	Indels 1	Gaps 1
Qy	1	MOTSFHKPFLLSMILAYSCCSLNGGGVAAETMVPQGIYDGETLTVSPRYTVIGDPSGTTVP	60	
Db	1	MOTSFHKPFLLSMILAYSCCSLNGGGVAAETMVPQGIYDGETLTVSPRYTVIGDPSGTTVP	60	
Qy	61	SAGELTLKLNLDNSIALPLSCFCPNLLGSPFVLGRGHSLLPENIRTSNCGAALSNSAADGL	120	
Db	61	SAGELTLKLNLDNSIALPLSCFCPNLLGSPFVLGRGHSLLPENIRTSNCGAALSNSAADGL	120	
Qy	121	FTIEGPKELSFSCNCSILLAVLPAAATNKGSGTPTTSTPSNGTIYSKTDLLLLLNKPKFSF	180	
Db	121	FTIEGPKELSFSCNCSILLAVLPAAATNKGSGTPTTSTPSNGTIYSKTDLLLLLNKPKFSF	180	
Qy	181	YSNLVSGDGAIDAKSITWQGIKLCVFOSENTAAQDGCQVTVTSFMANEAPIAFVAN	240	
Db	181	YSNLVSGDGAIDAKSITWQGIKLCVFOSENTAAQDGCQVTVTSFMANEAPIAFVAN	240	
Qy	241	VAGVRGGGIAAVDQGGQVSSSTSTEDPVVSFSRNTAVFDGNVARVGGGIYSYGNVAPL	300	
Db	241	VAGVRGGGIAAVDQGGQVSSSTSTEDPVVSFSRNTAVFDGNVARVGGGIYSYGNVAPL	300	
Qy	301	NNGKTLFLANNVAPSVYIAAKPQTSQGAQSNISNNYIGDGAIFCKNGAQ-AGSNNSGSVSPD	359	
Db	301	NNGKTLFLANNVAPSVYIAAQPTNNGQASNTSDNYIGDGAIFCKNGAQAGSNNSGSVSPD	360	
Qy	360	GBGVVPSSNVAAGKGAIYAKKLSVANGCPVQFLNIANDGCAIYLGESGELSLADYG	419	
Db	361	GBGVVPSSNVAAGKGAIYAKKLSVANGCPVQFLNIANDGCAIYLGESGELSLADYG	420	
Qy	420	DIIPDGNLKRATKENAADVNGVTVSSQAISMGSGGKIITLRKAKAGHQILFNDPIEMANGN	479	
Db	421	DIIPDGNLKRATKENAADVNGVTVSSQAISMGSGGKIITLRKAKAGHQILFNDPIEMANGN	480	
Qy	480	NQPAQSKLLKINDGEGYTGDIYFANGSFTLYQNVITIEQRIIVLRKAKLSVNSLSQTGG	539	
Db	481	NQPAQSKLLKINDGEGYTGDIYFANGSFTLYQNVITIEQRIIVLRKAKLSVNSLSQTGG	540	
Qy	540	SLYWEAGSTWDFVTPOPPQPPAANQILITLSNLHLSISLLANNVTPNPTNPQAODSHP	599	
Db	541	SLYWEAGSTLDFVTPOPPQPPAANQILITLSNLHLSISLLANNVTPNPTNPQAODSHP	600	
Qy	600	AVIGSTTAGSVTISGPIFFEDLDTDYDRYDMLGSGNQKINVLKQLGKTGPPANAPSDTL	659	
Db	601	AVIGSTTAGSVTISGPIFFEDLDTDYDRYDMLGSGNQKIDVLKQLGKTQPSANAPSDTL	660	
Qy	660	GNEMPKYGYGSKWLAWDPNTANGPYTLKATWTKGYNPGPERVASLVPNSLWGSILDI	719	
Db	661	GNBMPKYGYGSKWLAWDPNTANGPYTLKATWTKGYNPGPERVASLVPNSLWGSILDI	720	
Qy	720	RSASHTAQASVDGRSVCRLGVWSGVSNFFFYHDRDALDQGYRYTISGGVSYLGANSYFGSSMF	779	



Db 721 RSAHSAIQASVDGRSYCRGLVWSVNFYHDDALGQGYRYSIGSYLSGANSYFGSSMF 780  
QY 780 GLAFTVEFGRSKDYVVCRSNNHACIGSVYLSVSTQOALCGSYLPGDAPIRASVFGNQHMKT 839  
Db 781 GLAFTVEFGRSKDYVVCRSNNHACIGSVYLSVSTQOALCGSYLPGDAPIRASVFGNQHMKT 840  
QY 840 SYTFAEESDVRDNNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFVSADHESFTEEGD 899  
Db 841 SYTFAEESDVRDNNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFVSADHESFTEEGD 900  
QY 900 QARAFKSGHLLMLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959  
Db 901 QARAFKSGHLLMLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960  
QY 960 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYERDASRGYGLSAGSRVRF 1012  
Db 961 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYERDASRGYGLSAGSKVRF 1013

RESULT 5  
ADD43718  
ID ADD43718 standard; protein; 1013 AA.  
XX AC  
XX ADD43718;  
XX DT 15-JAN-2004 (first entry)  
XX DE Chlamydia trachomatis immunogenic protein, SEQ ID No 13.  
XX KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
XX KW gene therapy; antibacterial.  
XX OS Chlamydia trachomatis.  
XX PN W02003049762-A2.  
XX PD 19-JUN-2003.  
XX PF 12-DEC-2002; 2002WO-1B005761.  
XX PR 12-DEC-2001; 2001GB-00029732.  
XX PR 06-AUG-2002; 2002GB-00018233.  
XX PR 14-AUG-2002; 2002GB-00018924.  
XX (CHIR-) CHIRON SPA.  
XX PA Grandi G, Ratti G;  
XX PI WPI; 2003-532882/50.  
XX DR N-PSDB; ADD43719.  
XX PT New immunogenic composition having a protein or encoding nucleic acid,  
XX PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
XX PT infection.

XX Claim 6; SEQ ID NO 13; 164pp; English.  
XX CC The invention relates to a novel immunogenic composition comprising a  
XX CC protein or nucleic acid, and an adjuvant, where the protein or nucleic  
XX CC acid comprises any of 131 fully defined amino acid or nucleotide  
XX CC sequences given in the specification, or has 50% or greater sequence  
XX CC identity to it, or their fragments. The protein and/or nucleic acid of  
XX CC the immunogenic composition is useful in the manufacture of a medicament  
XX CC for the treatment or prevention of infection due to Chlamydia  
XX CC trachomatis. The infection is treated or prevented by the medicament  
XX CC eliciting an immune response which is specific to a C. trachomatis  
XX CC elementary body, or for neutralising C. trachomatis elementary bodies,  
XX CC hence the immunogenic composition can be used in creating a vaccine. The  
XX CC immunogenic compositions can also be used for the diagnosis of C.  
XX CC trachomatis infection. The nucleic acids of the immunogenic compositions  
XX CC can be used to treat disorders by gene therapy. The immunogenic  
XX CC compositions have antibacterial activity. This sequence represents one of

CC the 131 C. trachomatis proteins with immunogenic properties of the  
CC invention.  
XX SQ Sequence 1013 AA;  
Query Match 97.4%; Score 5131.5; DB 7; Length 1013;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;  
QY 1 MQTSFHKFELSMILAYSCSLNGGGYAAEIMVPMQGIYDGETLTVSPYTVIGDPSGTTVP 60  
Db 1 MQTSFHKFELSMILAYSCSLNGGGYAAEIMVPMQGIYDGETLTVSPYTVIGDPSGTTVP 60  
QY 61 SAGELTLKLNLDNSIAALPLSCFGLGSLTFLGRHSITFENIRTSITNGAALSNAADGL 120  
Db 61 SAGELTLKLNLDNSIAALPLSCFGLGSLTFLGRHSITFENIRTSITNGAALSNAADGL 120  
QY 121 FTIEGFKELSPNCNLSLAVLPAATTKNGSQTPPTTSPNGTIYKSTDLILLNNEKFSF 180  
Db 121 FTIEGFKELSPNCNLSLAVLPAATTKNGSQTPPTTSPNGTIYKSTDLILLNNEKFSF 180  
QY 181 YSNLVSGDGGAIIDAKSLTVQGISKLCVFQENTAQADGACOVVTSFSAMANEAPAFVAN 240  
Db 181 YSNLVSGDGGAIIDAKSLTVQGISKLCVFQENTAQADGACOVVTSFSAMANEAPAFVAN 240  
QY 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVEFDGNVAVRGVGIYSYGNVAPL 300  
Db 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVEFDGNVAVRGVGIYSYGNVAPL 300  
QY 301 NNGKTLFLNNVASPVYIAAKQPTSCQASNTSNYDGGAIKCNKGAQ-AGSNNSGSVSD 359  
Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAIKCNKGAQ-AGSNNSGSVSD 360  
QY 360 GEGVVFSSNVAAGKGAIYAKKLSVANGCPVQFLNINDGAIYLGESGELSADYVG 419  
Db 361 GEGVVFSSNVAAGKGAIYAKKLSVANGCPVQFLNINDGAIYLGESGELSADYVG 420  
QY 420 DIIFDGNLKRITAKENAADVGVTVSSQAI SMGSGGKIITLRAKAGHQLIFNDPIEMANGN 479  
Db 421 DIIFDGNLKRITAKENAADVGVTVSSQAI SMGSGGKIITLRAKAGHQLIFNDPIEMANGN 480  
QY 480 NQPAQSSKLLKINDGEYTGDIIVFANGSSTLYQNTVIEQRIVLREKAKLSVNSLSQTGG 539  
Db 481 NQPAQSSKLLKINDGEYTGDIIVFANGSSTLYQNTVIEQRIVLREKAKLSVNSLSQTGG 540  
QY 540 SLYMEAGSTWDFVTPQPPQPPAAANQLITLSNLHLSLSLLANNAVNTPTPPAQDSHP 599  
Db 541 SLYMEAGSTWDFVTPQPPQPPAAANQLITLSNLHLSLSLLANNAVNTPTPPAQDSHP 600  
QY 600 AVIGSTTAGSVTISGPIPPEDLDDTAYDRYDMLGSNQNINVLKQLGKTPPANAPSDTL 659  
Db 601 AVIGSTTAGSVTISGPIPPEDLDDTAYDRYDMLGSNQNINVLKQLGKTPPANAPSDTL 660  
QY 660 GNEMPKYGGQSKWLAWDPNTANNNGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDI 719  
Db 661 GNEMPKYGGQSKWLAWDPNTANNNGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDI 720  
QY 720 RSAHSAIQASVDGRSYCRGLVWSVNFYHDDALGQGYRYSIGSYLSGANSYFGSSMF 779  
Db 721 RSAHSAIQASVDGRSYCRGLVWSVNFYHDDALGQGYRYSIGSYLSGANSYFGSSMF 780  
QY 780 GLAFTVEFGRSKDYVVCRSNNHACIGSVYLSVSTQOALCGSYLPGDAPIRASVFGNQHMKT 839  
Db 781 GLAFTVEFGRSKDYVVCRSNNHACIGSVYLSVSTQOALCGSYLPGDAPIRASVFGNQHMKT 840  
QY 840 SYTFAEESDVRDNNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFVSADHESFTEEGD 899  
Db 841 SYTFAEESDVRDNNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFVSADHESFTEEGD 900  
QY 900 QARAFKSGHLLMLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959  
Db 901 QARAFKSGHLLMLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960



KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
KW gene therapy; antibacterial.  
XX Chlamydia trachomatis.  
OS WO2003049762-A2.  
XX 19-JUN-2003.  
PD 12-DEC-2002; 2002WO-IB005761.  
XX 12-DEC-2001; 2001GB-00029732.  
XX 06-AUG-2002; 2002GB-00018233.  
PR 14-AUG-2002; 2002GB-00018924.  
XX (CHIR-) CHIRON SPA.  
PA Grandi G, Ratti G;  
XX WPI; 2003-532882/50.  
XX N-PSDB; ADD43715.  
DR New immunogenic composition having a protein or encoding nucleic acid,  
XX useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
PT infection.  
PT Claim 6; SEQ ID NO 9; 164pp; English.  
XX The invention relates to a novel immunogenic composition comprising a  
XX protein or nucleic acid, and an adjuvant, where the protein or nucleic  
CC acid comprises any of 131 fully defined amino acid or nucleotide  
CC sequences given in the specification, or has 50% or greater sequence  
CC identity to it, or their fragments. The protein and/or nucleic acid of  
CC the immunogenic composition is useful in the manufacture of a medicament  
CC for the treatment or prevention of infection due to Chlamydia  
CC trachomatis. The infection is treated or prevented by the medicament  
CC eliciting an immune response which is specific to a C. trachomatis  
CC elementary body, or for neutralising C. trachomatis elementary bodies,  
CC hence the immunogenic composition can be used in creating a vaccine. The  
CC immunogenic compositions can also be used for the diagnosis of C.  
CC trachomatis infection. The nucleic acids of the immunogenic compositions  
CC can be used to treat disorders by gene therapy. The immunogenic  
CC compositions have antibacterial activity. This sequence represents one of  
CC the 131 C. trachomatis proteins with immunogenic properties of the  
CC invention.  
XX Sequence 1013 AA;  
SQ

Query Match 97.4%; Score 5131.5; DB 7; Length 1013;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 983; Conservative 19; Mismatches 11; Indels 1; Gaps 1;  
QY 1 MOTSFHKFLLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFYTIVGDPSTTVF 60  
DB 1 MOTSFHKFLLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFYTIVGDPSTTVF 60  
QY 61 SAGELTLKNDNSIAALPLSCFNLGSPFTVLGRGHSHTPENITRTSTNGAALSNSAADGL 120  
DB 61 SAGELTLKNDNSIAALPLSCFNLGSPFTVLGRGHSHTPENITRTSTNGAALSNSAADGL 120  
QY 121 FTTEGFKELSPNSCNLLAVLPAATNKGSGTPTTTPSNGTIYSKTDILLNNEKFSF 180  
DB 121 FTTEGFKELSPNSCNLLAVLPAATNKGSGTPTTTPSNGTIYSKTDILLNNEKFSF 180  
QY 181 YSNLVSGDGAIDAKSLTVQGISLKVFOENTAGDGAQCVVTSFSAMANEAPAFVAN 240  
DB 181 YSNLVSGDGAIDAKSLTVQGISLKVFOENTAGDGAQCVVTSFSAMANEAPAFVAN 240  
QY 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVFSFSNTAVFEDGNVARYGGGIYSYGNVAF 300  
DB 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVFSFSNTAVFEDGNVARYGGGIYSYGNVAF 300  
QY 301 NNGKTLFLNNVASFVYIAAEQPTNGQASNTSDNYTGGAIFCKGQAAGSNNSGSVDF 359

DB 301 NNGKTLFLNNVASFVYIAAEQPTNGQASNTSDNYTGGAIFCKGQAAGSNNSGSVDF 360  
QY 360 GEGVVFSSNVAAGKGGAIYAKLVANCGPVQFQFVFNRIANDGGAIIYLGESGELSADYG 419  
DB 361 GEGVVFSSNVAAGKGGAIYAKLVANCGPVQFQFVFNRIANDGGAIIYLGESGELSADYG 420  
QY 420 DIIFDGNLKRITAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLFNDPIEMANGN 479  
DB 421 DIIFDGNLKRITAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLFNDPIEMANGN 480  
QY 480 NQPAOSSKLLKINDGEGYTGDIIVFANGSSTLYQNTYIEQGRIVLREKAKLVNSLSQSGG 539  
DB 481 NQPAOSSKLLKINDGEGYTGDIIVFANGSSTLYQNTYIEQGRIVLREKAKLVNSLSQSGG 540  
QY 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPPTPPAQDSDP 599  
DB 541 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPPTPPAQDSDP 600  
QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQKINVLKQLQGTQPSANAPSDLT 659  
DB 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQKIDVLKQLQGTQPSANAPSDLT 660  
QY 660 GNEMPKYQGSWKLAWDPNTANNPGPYTLKATWTKTYNPGPERSVSLVNSIWSGILDI 719  
DB 661 GNEMPKYQGSWKLAWDPNTANNPGPYTLKATWTKTYNPGPERSVSLVNSIWSGILDI 720  
QY 720 RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDALGQGYRYISGGYSLGANSYFGSSMF 779  
DB 721 RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDALGQGYRYISGGYSLGANSYFGSSMF 780  
QY 780 GLAFTVEFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPFIRASYGFGNQHMT 839  
DB 781 GLAFTVEFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPFIRASYGFGNQHMT 840  
QY 840 SYTPAESDVRWNNCLAGEIGAGLPVITPDKLYNELRPFVQAEFSYADHESFTEEGD 899  
DB 841 SYTPAESDVRWNNCLAGEIGAGLPVITPDKLYNELRPFVQAEFSYADHESFTEEGD 900  
QY 900 QARAFKSHLLNLVSPGVKFCRSTHBNKYSFMAAYICDAVRTISGTTETLLSHQETW 959  
DB 901 QARAFKSHLLNLVSPGVKFCRSTHBNKYSFMAAYICDAVRTISGTTETLLSHQETW 960  
QY 960 TTDAFLARHGIVVVRGSMYASLTSTNIEVYGHGRYEDRSDASRGYGLSAGSRVRF 1012  
DB 961 TTDAFLARHGIVVVRGSMYASLTSTNIEVYGHGRYEDRSDASRGYGLSAGSRVRF 1013

RESULT 8  
ADD43786  
ID ADD43786 standard; protein; 1013 AA.  
XX  
AC ADD43786;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Chlamydia trachomatis immunogenic protein, SEQ ID No 81.  
XX  
KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;  
KW gene therapy; antibacterial.  
XX Chlamydia trachomatis.  
OS  
XX  
PN WO2003049762-A2.  
XX  
PD 19-JUN-2003.  
XX  
PF 12-DEC-2002; 2002WO-IB005761.  
XX  
PR 12-DEC-2001; 2001GB-00029732.  
PR 06-AUG-2002; 2002GB-00018233.  
PR 14-AUG-2002; 2002GB-00018924.  
XX

PA	(CHIR-) CHIRON SPA.	
XX		
PI	Grandi G, Ratti G;	
XX		
DR	WPI; 2003-532882/50.	
XX	N-PSDB; ADD43787.	
XX		
PT	New immunogenic composition having a protein or encoding nucleic acid,	
PT	useful for diagnosing, preventing and/or treating Chlamydia trachomatis	
PT	infection.	
XX		
PS	Claim 6; SEQ ID NO 81; 164pp; English.	
XX		
CC	The invention relates to a novel immunogenic composition comprising a	
CC	protein or nucleic acid, and an adjuvant, where the protein or nucleic	
CC	acid comprises any of 131 fully defined amino acid or nucleotide	
CC	sequences given in the specification, or has 50% or greater sequence	
CC	identity to it, or their fragments. The protein and/or nucleic acid of	
CC	the immunogenic composition is useful in the manufacture of a medicament	
CC	for the treatment or prevention of infection due to Chlamydia	
CC	trachomatis. The infection is treated or prevented by the medicament	
CC	eliciting an immune response which is specific to a C. trachomatis	
CC	elementary body, or for neutralising C. trachomatis elementary bodies,	
CC	hence the immunogenic composition can be used in creating a vaccine. The	
CC	immunogenic compositions can also be used for the diagnosis of C.	
CC	trachomatis infection. The nucleic acids of the immunogenic compositions	
CC	can be used to treat disorders by gene therapy. The immunogenic	
CC	compositions have antibacterial activity. This sequence represents one of	
CC	the 131 C. trachomatis proteins with immunogenic properties of the	
CC	invention.	
XX		
XX	Sequence 1013 AA;	

Query Match	97.4%; Score 5131.5; DB 7; Length 1013;	
Best Local Similarity	97.0%; Pred. No. 0;	
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;		
QY	1 MQTSFHKPFLSMILAYSCSLNGGYYAAEIMVPGIYDGETLTVSPFTVTVGDSGTTVF	60
DB	1 MQTSFHKPFLSMILAYSCSLNGGYYAAEIMVPGIYDGETLTVSPFTVTVGDSGTTVF	60
QY	61 SAGELTLKNLNSIAALPLSCFNLGSLTFVLGRHSITFENIRTSNGAALSAAADGL	120
DB	61 SAGELTLKNLNSIAALPLSCFNLGSLTFVLGRHSITFENIRTSNGAALSAAADGL	120
QY	121 FTIEGPKELSPSNCNLSLAVLPAATNNKSGSTPTTTPSNGTIYSKTDLLLNKPSF	180
DB	121 FTIEGPKELSPSNCNLSLAVLPAATNNKSGSTPTTTPSNGTIYSKTDLLLNKPSF	180
QY	181 YSNLVSGDGGDAIDAKSLTVQISKLCVPOENTAQADGGACQVTSFSMAANEAPFAN	240
DB	181 YSNLVSGDGGDAIDAKSLTVQISKLCVPOENTAQADGGACQVTSFSMAANEAPFAN	240
QY	241 VAGVGGGIAAVDQGGQVSSSTSTEDPVVSFRNTAVEPDGNVARGGGIYSYGNVAF	300
DB	241 VAGVGGGIAAVDQGGQVSSSTSTEDPVVSFRNTAVEPDGNVARGGGIYSYGNVAF	300
QY	301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAIKCKGAQAGSNNSGVSFD	359
DB	301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAIKCKGAQAGSNNSGVSFD	360
QY	360 GEGVFFPSNVAAGKGGAIYAKLSVANCGPVQFLRNITANDGGAIIYGESGELSADYG	419
DB	361 GEGVFFPSNVAAGKGGAIYAKLSVANCGPVQFLRNITANDGGAIIYGESGELSADYG	420
QY	420 DIIFDGNLRTAKENAAVDVGVYSSQAI SMGSGKITTLAKAGHQILFNDPIEMANGN	479
DB	421 DIIFDGNLRTAKENAAVDVGVYSSQAI SMGSGKITTLAKAGHQILFNDPIEMANGN	480
QY	480 NPAQSSKLLKINDGEGTGDIVFANGSSTLYQNVITBQGRIVLREKAKLSVNSLSQTGG	539
DB	481 NPAQSSKLLKINDGEGTGDIVFANGSSTLYQNVITBQGRIVLREKAKLSVNSLSQTGG	540

QY	540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNVNTPTNPPAQDSHP	599
DB	541 SLYMEAGSTLDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNVNTPTNPPAQDSHP	600
QY	600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKNQKINVLKQLGTGKPPANAPSDLTL	659
DB	601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKNQKIDVLKQLGTGKPPANAPSDLTL	660
QY	660 GNEMPKYGYQGSWKLAWDPNTANNPVTTLKATWTKTYNPGPERSVASLVPNSLWGSILDI	719
DB	661 GNEMPKYGYQGSWKLAWDPNTANNPVTTLKATWTKTYNPGPERSVASLVPNSLWGSILDI	720
QY	720 RSAHSAIOASVDGSSYCRGLWVSGVSNPFYHRRDALGQGYRISGGYSLGANSYFGSSMF	779
DB	721 RSAHSAIOASVDGSSYCRGLWVSGVSNPFYHRRDALGQGYRISGGYSLGANSYFGSSMF	780
QY	780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQALCGSVLFGDAFIRASVGFQGNHMK	839
DB	781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQALCGSVLFGDAFIRASVGFQGNHMK	840
QY	840 SYTPAESDVRWDMNCLAGSIGAGLPITVITPSKLYLNLRLPFFVQAEFSYADHESFTTEGD	899
DB	841 SYTPAESDVRWDMNCLAGSIGAGLPITVITPSKLYLNLRLPFFVQAEFSYADHESFTTEGD	900
QY	900 QARAPKSHLLNLSPVGVKFXDRCSSTHPNKYSFMAAYICDAYRTISGTETTLISHQETW	959
DB	901 QARAPKSHLLNLSPVGVKFXDRCSSTHPNKYSFMAAYICDAYRTISGTETTLISHQETW	960
QY	960 TTDAFLHARHGVVVRGSMYASLTNSIEVYGHGRYEDRSDASRGYGLSAGSRVRF	1012
DB	961 TTDAFLHARHGVVVRGSMYASLTNSIEVYGHGRYEDRSDASRGYGLSAGSRVRF	1013
RESULT 9		
ADD43788		
ID	ADD43788 standard; protein; 1013 AA.	
XX		
AC	ADD43788;	
XX		
DT	15-JAN-2004 (first entry)	
XX		
DE	Chlamydia trachomatis immunogenic protein, SEQ ID NO 83.	
XX		
KW	immunogenic; infection; Chlamydia trachomatis; immune; vaccine;	
XX	gene therapy; antibacterial.	
OS	Chlamydia trachomatis.	
XX		
PN	WO2003049762-A2.	
XX		
PD	19-JUN-2003.	
XX		
PF	12-DEC-2002; 2002WO-IB005761.	
XX		
PR	12-DEC-2001; 2001GB-00029732.	
PR	06-AUG-2002; 2002GB-00018233.	
PR	14-AUG-2002; 2002GB-00018924.	
XX		
PA	(CHIR-) CHIRON SPA.	
XX		
PI	Grandi G, Ratti G;	
XX		
DR	WPI; 2003-532882/50.	
DR	N-PSDB; ADD43789.	
XX		
PT	New immunogenic composition having a protein or encoding nucleic acid,	
PT	useful for diagnosing, preventing and/or treating Chlamydia trachomatis	
PT	infection.	
XX		
PS	Claim 6; SEQ ID NO 83; 164pp; English.	
XX		
CC	The invention relates to a novel immunogenic composition comprising a	
CC	protein or nucleic acid, and an adjuvant, where the protein or nucleic	

CC acid comprises any of 131 fully defined amino acid or nucleotide  
CC sequences given in the specification, or has 50% or greater sequence  
CC identity to it, or their fragments. The protein and/or nucleic acid of  
CC the immunogenic composition is useful in the manufacture of a medicament  
CC for the treatment or prevention of infection due to Chlamydia  
CC trachomatis. The infection is treated or prevented by the medicament  
CC eliciting an immune response which is specific to a C. trachomatis  
CC elementary body, or for neutralising C. trachomatis elementary bodies,  
CC hence the immunogenic composition can be used in creating a vaccine. The  
CC immunogenic compositions can also be used for the diagnosis of C.  
CC trachomatis infection. The nucleic acids of the immunogenic compositions  
CC can be used to treat disorders by gene therapy. The immunogenic  
CC compositions have antibacterial activity. This sequence represents one of  
CC the 131 C. trachomatis proteins with immunogenic properties of the  
CC invention.

XX  
SQ Sequence 1013 AA;

Query Match		97.4%;	Score 5131.5;	DB 7;	Length 1013;
Best Local Similarity		97.0%;	Pred. No. 0;		
Matches 983;		Conservative 18;	Mismatches 11;	Indels 1;	Gaps 1;
Qy	1	MQTSFHKPFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPGGTVF	60		
Db	1	MQTSFHKPFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPGGTVF	60		
Qy	61	SAGELTLKMLNSIAALPLSCFNLGSLTFTVGRHSLTFENIRSTNGAALSNAADGL	120		
Db	61	SAGELTLKMLNSIAALPLSCFNLGSLTFTVGRHSLTFENIRSTNGAALSNAADGL	120		
Qy	121	FTIEGPKELSPNCNSLLAVLPAATNKGSTQPTTSTPSNGTIYSKTDLLLNNEKFSF	180		
Db	121	FTIEGPKELSPNCNSLLAVLPAATNKGSTQPTTSTPSNGTIYSKTDLLLNNEKFSF	180		
Qy	181	YSLNLSVGGGGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTSPSAMANEAPAFVAN	240		
Db	181	YSLNLSVGGGGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTSPSAMANEAPAFVAN	240		
Qy	241	VAGVRGGGIAAVQDQGGQVSSSTSTEDPVVPSRNTAVBFDGNVAVRGGGIYSYGNVAF	300		
Db	241	VAGVRGGGIAAVQDQGGQVSSSTSTEDPVVPSRNTAVBFDGNVAVRGGGIYSYGNVAF	300		
Qy	301	NGKTLFLNNVSPVYIAAKQPTSGOASNTSNYGDGGAIFCKNGAQ-AGSNNSGVSFD	359		
Db	301	NGKTLFLNNVSPVYIAAEQPTNQASNTSDNYGDGGAIFCKNGAQAGSNNSGVSFD	360		
Qy	360	GEGVWPFSSNVAAGKGAIVAKKLSVANGCPVQFLRNIAANDGGAIVLGBSGELSLSDIYG	419		
Db	361	GEGVWPFSSNVAAGKGAIVAKKLSVANGCPVQFLRNIAANDGGAIVLGBSGELSLSDIYG	420		
Qy	420	DIIFDGNLKRITAKENAADVNGVTVSSQAISMGGGKIITLRKAKVHQLIFNDPIEWANGN	479		
Db	421	DIIFDGNLKRITAKENAADVNGVTVSSQAISMGGGKIITLRKAKVHQLIFNDPIEWANGN	480		
Qy	480	NQPAQSSKLLKINDGEGYTDIIPANGSSTLYQNVTEIQGRIVLRKAKLSVNSLSQTGG	539		
Db	481	NQPAQSSKLLKINDGEGYTDIIPANGSSTLYQNVTEIQGRIVLRKAKLSVNSLSQTGG	540		
Qy	540	SLYWEAGSTWDFVTPPPQPPQPPAANOLITLSNLHLSLSLLANNAVTPNTPPAQDSHP	599		
Db	541	SLYWEAGSTWDFVTPPPQPPQPPAANOLITLSNLHLSLSLLANNAVTPNTPPAQDSHP	600		
Qy	600	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDVLGWSNQKINVLKQLGKTPPANASDLTL	659		
Db	601	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDVLGWSNQKIDVLKQLGKTPPANASDLTL	660		
Qy	660	GNEMPKYQGSWKLAWDPNTANNPPTLKATWTKGTGYNPGERVASLVPNSLWGSFLDI	719		
Db	661	GNEMPKYQGSWKLAWDPNTANNPPTLKATWTKGTGYNPGERVASLVPNSLWGSFLDI	720		
Qy	720	RSASAIQASVDCGRSYCRGLVWGSVSNFFVHRRDALGQGVRYISGGYSLGANSYFGSSMF	779		
Db	721	RSASAIQASVDCGRSYCRGLVWGSVSNFFVHRRDALGQGVRYISGGYSLGANSYFGSSMF	780		

Qy	780	GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLTOQALCGSYLFGDAFTRASYGFGNQHMKT	839
Db	781	GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLTKQALCGSYLFGDAFTRASYGFGNQHMKT	840
Qy	840	SYTFAESDVWDNNCLAGEIGAGLPVITPSPKLYLNEIRPFVQAEFSYADHESFTBERGD	899
Db	841	SYTFAESDVWDNNCLVGEIGVGLPVIPTPSKLYLNEIRPFVQAEFSYADHESFTBERGD	900
Qy	900	QARAFKSHLNLNLSPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	959
Db	901	QARAFKSHLNLNLSPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	960
Qy	960	TTDAFHLARHGCVVVRGSGMYASLTSNIEVYHGRYEYRDASRGYGLSAGSRVRF	1012
Db	961	TTDAFHLARHGCVVVRGSGMYASLTSNIEVYHGRYEYRDTSRGYGLSAGSKVRF	1013

RESULT 10

ADD43710

ID ADD43710 standard; protein; 1013 AA.

AC ADD43710;

DT 15-JAN-2004 (first entry)

DE Chlamydia trachomatis immunogenic protein, SEQ ID No 5.

EE immunogenic; infection; Chlamydia trachomatis; immune; vaccine;

KW gene therapy; antibacterial.

OS Chlamydia trachomatis.

PN WO2003049762-A2.

PD 19-JUN-2003.

PF 12-DEC-2002; 2002WO-IB005761.

PR 12-DEC-2001; 2001GB-00029732.

PR 06-AUG-2002; 2002GB-00018233.

PR 14-AUG-2002; 2002GB-00018924.

XX (CHIR-) CHIRON SPA.

XX Grandi G, Ratti G;

XX WPI; 2003-532882/50.

DR N-PSDB; ADD43711.

XX New immunogenic composition having a protein or encoding nucleic acid,  
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis  
PT infection.

XX Claim 6; SEQ ID NO 5; 164pp; English.

XX The invention relates to a novel immunogenic composition comprising a  
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic  
CC acid comprises any of 131 fully defined amino acid or nucleotide  
CC sequences given in the specification, or has 50% or greater sequence  
CC identity to it, or their fragments. The protein and/or nucleic acid of  
CC the immunogenic composition is useful in the manufacture of a medicament  
CC for the treatment or prevention of infection due to Chlamydia  
CC trachomatis. The infection is treated or prevented by the medicament  
CC eliciting an immune response which is specific to a C. trachomatis  
CC elementary body, or for neutralising C. trachomatis elementary bodies,  
CC hence the immunogenic composition can be used in creating a vaccine. The  
CC immunogenic compositions can also be used for the diagnosis of C.  
CC trachomatis infection. The nucleic acids of the immunogenic compositions  
CC can be used to treat disorders by gene therapy. The immunogenic  
CC compositions have antibacterial activity. This sequence represents one of  
CC the 131 C. trachomatis proteins with immunogenic properties of the  
CC invention.



QY 121 FTIEGKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180  
 DB 121 FTIEGKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180  
 QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTSPSAMANEAPAFVAN 240  
 DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTSPSAMANEAPAFVAN 240  
 QY 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFEDGNVAVGGGIYSYGNVAF 300  
 DB 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFEDGNVAVGGGIYSYGNVAF 300  
 QY 301 NNGKTLFLNNVSPVYIAAKOPTSGQASNTSNNGYDGGAI FCKNGAQ-AGSNNSGSVSD 359  
 DB 301 NNGKTLFLNNVSPVYIAAKOPTSGQASNTSNNGYDGGAI FCKNGAQ-AGSNNSGSVSD 359  
 QY 360 GEGVVFSSNVAAGKGAIVAKLSVANGCPVQFLNANIANDGGAIYLGSGELSLADYG 419  
 DB 361 GEGVVFSSNVAAGKGAIVAKLSVANGCPVQFLNANIANDGGAIYLGSGELSLADYG 420  
 QY 420 DIIFDGNLKRKTAKENADVNGVTVSSQAISMGSGKITTIRAKAGHOIILFNDPIEMANG 479  
 DB 421 DIIFDGNLKRKTAKENADVNGVTVSSQAISMGSGKITTIRAKAGHOIILFNDPIEMANG 480  
 QY 480 NPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTEIOGRI VLREKAKLSVNSLSQTGG 539  
 DB 481 NPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNVTEIOGRI VLREKAKLSVNSLSQTGG 540  
 QY 540 SLVMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSLLANNVTPNPPPAQDSHP 599  
 DB 541 SLVMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSLLANNVTPNPPPAQDSHP 600  
 QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKINVLKQLGTQPPANAPSDTL 659  
 DB 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKINVLKQLGTQPPANAPSDTL 660  
 QY 660 GNEMPKYGVGSKLAWDPNTANGPYTLKATWTKYGNPGERVASLVPNSLWGSTLDI 719  
 DB 661 GNEMPKYGVGSKLAWDPNTANGPYTLKATWTKYGNPGERVASLVPNSLWGSTLDI 720  
 QY 720 RSAHSAIQASVDCGRSYCRGLWVGVSNNFFYHDDALGQGYRIYSGGYSLGANSYFGSSMF 779  
 DB 721 RSAHSAIQASVDCGRSYCRGLWVGVSNNFFYHDDALGQGYRIYSGGYSLGANSYFGSSMF 780  
 QY 780 GLAFTEVFGSKDYVVCRSNHHACIGSVYLSSTQALCGSVLFGDAFIRASYGFGNQHMT 839  
 DB 781 GLAFTEVFGSKDYVVCRSNHHACIGSVYLSSTQALCGSVLFGDAFIRASYGFGNQHMT 840  
 QY 840 SYTFAESDVRWNNCLAGEIGAGLPVITPSKLYLNLRLPFOQAEFSYADHESFTEGD 899  
 DB 841 SYTFAESDVRWNNCLAGEIGAGLPVITPSKLYLNLRLPFOQAEFSYADHESFTEGD 900  
 QY 900 QARAFKSHLLNLSPVGVKFDRCSTHPNKSYPMAAYICDAVRTISGTETTLISHOETW 959  
 DB 901 QARAFKSHLLNLSPVGVKFDRCSTHPNKSYPMAAYICDAVRTISGTETTLISHOETW 960  
 QY 960 TTDAFHLARHGVVVRGSMYASLTNSIEVYGHGRYEYRDAISRGYGLSAGSRVF 1012  
 DB 961 TTDAFHLARHGVVVRGSMYASLTNSIEVYGHGRYEYRDAISRGYGLSAGSRVF 1013

RESULT 12

ID ADW29027 standard; protein; 1013 AA.

XX AC ADW29027;

XX DT 07-APR-2005 (first entry)

XX DE C<sub>2</sub>trachomatis polymorphic membrane protein (PmpG) Seq 21.

XX KW immunogenicity; antigen; vaccine; chlamydia trachomatis infection; antibacterial; gene therapy.

XX Chlamydia trachomatis.  
 OS WO2005002619-A2.  
 XX PN 13-JAN-2005.  
 XX PD 25-JUN-2004; 2004WO-US020491.  
 XX PF 26-JUN-2003; 2003GB-00015020.  
 XX PR 25-AUG-2003; 2003US-0497649P.  
 XX PR 02-FEB-2004; 2004GB-00002236.  
 XX PR 01-JUN-2004; 2004US-0576375P.  
 XX PA (CHIR ) CHIRON CORP.  
 XX PI Grandi G, Finco O, Ratti G, Bonci A;  
 XX WPI; 2005-075653/08.  
 XX DR  
 XX PT An immunogenic composition for preventing or treating Chlamydia trachomatis infections comprises a combination of C. trachomatis antigens, such as PepA, LcrB, ArtJ, DnaK, CT398, OmpH-like, L7/L12, OmcA, or Atos.  
 XX PS Example 1; SEQ ID NO 21; 110pp; English.  
 XX CC This invention relates to novel immunogenic compositions that comprise a combination of Chlamydia trachomatis (C. trachomatis) antigens. Specifically, it refers to a first antigen group consisting of PepA, LcrB, ArtJ, DnaK and CT398 and a second antigen group consisting of PepA, LcrB, ArtJ, DnaK and CT398, OmpH-like, L7/L12, OmcA, Atos, CT547, Enolase, HtrA and MurG. The present invention describes the development of a vaccine and antibodies that recognize the immunogenic composition in order to provide a method of neutralizing a C. trachomatis infection in a mammal through raising an immune response. Accordingly, the pharmaceutical compositions derived thereof exhibit antibacterial activity and the group of antigenic DNA sequences can also be used in gene therapy. This polypeptide sequence is a Chlamydia trachomatis antigen of the invention.  
 XX SQ Sequence 1013 AA;  
 QY Query Match 97.4%; Score 5131.5; DB 9; Length 1013;  
 DB Best Local Similarity 97.0%; Pred. No. 0;  
 QY Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;  
 DB 1 MOTSFHKFPLSMILAYSCCSLNGGGYAAIMVPPQGIYDGETLTVSPPYTVIGDPSGTTVF 60  
 QY 1 MOTSFHKFPLSMILAYSCCSLNGGGYAAIMVPPQGIYDGETLTVSPPYTVIGDPSGTTVF 60  
 DB 61 SAGELTKNLDNSIAALPLSCFNLGSGFTVLGRGHSHTFENIRTSNKAALSNSAADGL 120  
 QY 61 SAGELTKNLDNSIAALPLSCFNLGSGFTVLGRGHSHTFENIRTSNKAALSNSAADGL 120  
 DB 121 FTIEGKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180  
 QY 121 FTIEGKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF 180  
 DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTSPSAMANEAPAFVAN 240  
 QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTSPSAMANEAPAFVAN 240  
 DB 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFEDGNVAVGGGIYSYGNVAF 300  
 QY 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFEDGNVAVGGGIYSYGNVAF 300  
 DB 301 NNGKTLFLNNVSPVYIAAKOPTSGQASNTSNNGYDGGAI FCKNGAQ-AGSNNSGSVSD 359  
 QY 301 NNGKTLFLNNVSPVYIAAKOPTSGQASNTSNNGYDGGAI FCKNGAQ-AGSNNSGSVSD 359  
 DB 360 GEGVVFSSNVAAGKGAIVAKLSVANGCPVQFLNANIANDGGAIYLGSGELSLADYG 419  
 QY 360 GEGVVFSSNVAAGKGAIVAKLSVANGCPVQFLNANIANDGGAIYLGSGELSLADYG 419



Db 361 GBGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQPLGNIANDDGAIYLGESGELSADYG 420  
Qy 420 DIIPGNLKRRTAKENAADVNGVTVSSQAISMGSGGKITTLRAKAGHOILFNDPIEMANGN 479  
Db 421 DIIPGNLKRRTAKENAADVNGVTVSSQAISMGSGGKITTLRAKAGHOILFNDPIEMANGN 480  
Qy 480 NQPAQSSKLLKINDGEGYTDIVFANGSTLYQNTVIEQGRIVLREKAKLSVNSLSQGG 539  
Db 481 NQPAQSSKLLKINDGEGYTDIVFANGSTLYQNTVIEQGRIVLREKAKLSVNSLSQGG 540  
Qy 540 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599  
Db 541 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 600  
Qy 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSKNINVLKQLGTPPANAPSDTL 659  
Db 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSKNINVLKQLGTPPANAPSDTL 660  
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNPVTYTKATWTGTGYNPGRVVASLVPNSLWGSILDI 719  
Db 661 GNEMPKYGYQGSWKLAWDPNTANNPVTYTKATWTGTGYNPGRVVASLVPNSLWGSILDI 720  
Qy 720 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDALGQGYRIYSGYSLGANSYFGSSWF 779  
Db 721 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDALGQGYRIYSGYSLGANSYFGSSWF 780  
Qy 780 GLAFTVEVGRSKDYVVCRSNHHACIGSVYLSFQOALCGSYLFGDAFIRASVYFGNQHMKT 839  
Db 781 GLAFTVEVGRSKDYVVCRSNHHACIGSVYLSFQOALCGSYLFGDAFIRASVYFGNQHMKT 840  
Qy 840 SVTPAEBSVDVRDNNCLAGIGAGLPIVITPESKLYLNELRPVQAEPSYADHESPTREGD 899  
Db 841 SVTPAEBSVDVRDNNCLAGIGAGLPIVITPESKLYLNELRPVQAEPSYADHESPTREGD 900  
Qy 900 QARAFKSGHLLNLSVPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETW 959  
Db 901 QARAFKSGHLLNLSVPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETW 960  
Qy 960 TTDAFHLARHGVVVRGSMYASLTSNIYEVGHGRYRYDRASRGYGLSAGSRVRF 1012  
Db 961 TTDAFHLARHGVVVRGSMYASLTSNIYEVGHGRYRYDRASRGYGLSAGSRVRF 1013

RESULT 13  
ABAL9078  
ID ABAL9078 standard; protein; 1013 AA.  
XX  
AC ABAL9078;  
XX  
DT 28-JUL-2005 (first entry)  
XX  
DE Chlamydia trachomatis protein - SEQ ID 81.  
XX  
KW chlamydia trachomatis infection; antibacterial; vaccine.  
XX  
OS Chlamydia trachomatis.  
XX  
PN US2005106162-A1.  
XX  
PD 19-MAY-2005.  
XX  
PF 16-DEC-2004; 2004US-00498327.  
XX  
PR 12-DEC-2001; 2001GB-00029732.  
XX  
PR 06-AUG-2002; 2002EP-00182233.  
XX  
PR 14-AUG-2002; 2001EP-00218924.  
XX  
PR 12-DEC-2002; 2002WO-IB005761.  
XX  
PA (GRAN/) GRANDI G.  
XX  
PA (RATTI/) RATTI G.  
XX  
PI Grandi G, Ratti G;  
XX

DR MPI; 2005-354777/36.  
DR N-FSDB; ABAL9079.  
XX  
PT New immunogenic composition comprising Chlamydia trachomatis protein or  
PT its encoding nucleic acid, useful for immunizing against, or treating or  
PT preventing chlamydial infection.  
XX  
PS Claim 13; SEQ ID NO 81; 90pp; English.  
XX  
CC The invention comprises an immunogenic composition for the treatment or  
CC prevention of a Chlamydia trachomatis infection. The immunogenic  
CC composition contains a Chlamydia trachomatis protein or its encoding  
CC nucleic acid. The immunogenic composition of the invention is useful  
CC treating or preventing chlamydial infection. The present amino acid  
CC sequence represents a Chlamydia trachomatis protein of the invention.  
CC NOTE: The present sequence is not shown in the specification, but was  
CC obtained from the USPTO website -  
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.  
XX  
SQ Sequence 1013 AA;  
Query Match 97.4%; Score 5131.5; DB 9; Length 1013;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;  
Qy 1 MOTSFHKFFLLSMILAYSCCSLNGGGYAAIMVPOGIYDGETLTVSPFVTVIGDPSGTTVF 60  
Db 1 MOTSFHKFFLLSMILAYSCCSLNGGGYAAIMVPOGIYDGETLTVSPFVTVIGDPSGTTVF 60  
Qy 61 SAGELTKNLNDNSTAALPLSCFNLGSPFVTLGRGHSITPENIRTSNTGAALSAAADGL 120  
Db 61 SAGELTKNLNDNSTAALPLSCFNLGSPFVTLGRGHSITPENIRTSNTGAALSAAADGL 120  
Qy 121 FTIEGFKELSPNCNLSLLAVLPAATTNKGSTPTTSTPSNGTITSKTDLLLANNEKPSF 180  
Db 121 FTIEGFKELSPNCNLSLLAVLPAATTNKGSTPTTSTPSNGTITSKTDLLLANNEKPSF 180  
Qy 181 YSNLVSDDGGAIDAKSLTVQGISKLCYFOENTAODGACOVVTSFSANAEAFVAFAN 240  
Db 181 YSNLVSDDGGAIDAKSLTVQGISKLCYFOENTAODGACOVVTSFSANAEAFVAFAN 240  
Qy 241 VAGVGGGIAAVODQOQGVSSSTSTEDPVVSFSNTAVEPDGNVAVGGGIYSYGNVAF 300  
Db 241 VAGVGGGIAAVODQOQGVSSSTSTEDPVVSFSNTAVEPDGNVAVGGGIYSYGNVAF 300  
Qy 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAAGSNNSGSVSD 359  
Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAAGSNNSGSVSD 360  
Qy 360 GBGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQPLGNIANDDGAIYLGESGELSADYG 419  
Db 361 GBGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQPLGNIANDDGAIYLGESGELSADYG 420  
Qy 420 DIIPGNLKRRTAKENAADVNGVTVSSQAISMGSGGKITTLRAKAGHOILFNDPIEMANGN 479  
Db 421 DIIPGNLKRRTAKENAADVNGVTVSSQAISMGSGGKITTLRAKAGHOILFNDPIEMANGN 480  
Qy 480 NQPAQSSKLLKINDGEGYTDIVFANGSTLYQNTVIEQGRIVLREKAKLSVNSLSQGG 539  
Db 481 NQPAQSSKLLKINDGEGYTDIVFANGSTLYQNTVIEQGRIVLREKAKLSVNSLSQGG 540  
Qy 540 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599  
Db 541 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 600  
Qy 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSKNINVLKQLGTPPANAPSDTL 659  
Db 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSKNINVLKQLGTPPANAPSDTL 660  
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNPVTYTKATWTGTGYNPGRVVASLVPNSLWGSILDI 719  
Db 661 GNEMPKYGYQGSWKLAWDPNTANNPVTYTKATWTGTGYNPGRVVASLVPNSLWGSILDI 720

QY 720 RSHSAIQASVDRGSRVCRGLWVSGVSNPFYHDDALGQGYRISGGYSIGANSYFGSSMF 779  
DB 721 RSHSAIQASVDRGSRVCRGLWVSGVSNPFYHDDALGQGYRISGGYSIGANSYFGSSMF 780  
QY 780 GLAFTEVFGSRKDYVVCVCRSNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMKT 839  
DB 781 GLAFTEVFGSRKDYVVCVCRSNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMKT 840  
QY 840 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFSADHESFTEGD 899  
DB 841 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFSADHESFTEGD 900  
QY 900 QARAFKSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 959  
DB 901 QARAFRSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 960  
QY 960 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDSRGYLSAGSKVRFP 1012  
DB 961 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDSRGYLSAGSKVRFP 1013

RESULT 14

AEA19002  
ID AEA19002 standard; protein; 1013 AA.  
AC AEA19002;

XX 28-JUL-2005 (first entry)

XX Chlamydia trachomatis protein - SEQ ID 5.

XX chlamydia trachomatis infection; antibacterial; vaccine.

XX Chlamydia trachomatis.

XX US2005106162-A1.

XX 19-MAY-2005.

XX 16-DEC-2004; 2004US-00498327.

XX 12-DEC-2001; 2001GB-00029732.

XX 06-AUG-2002; 2002EP-00182233.

XX 14-AUG-2002; 2001EP-00218924.

XX 12-DEC-2002; 2002WO-IB005761.

XX (GRAN/) GRANDI G.

XX (RATT/) RATTI G.

XX Grandi G, Ratti G;

XX WPI; 2005-354777/36.

XX N-PSDB; AEA19003.

XX New immunogenic composition comprising Chlamydia trachomatis protein or

XX its encoding nucleic acid, useful for immunizing against, or treating or

XX preventing chlamydial infection.

XX Claim 13; SEQ ID NO 5; 90pp; English.

XX The invention comprises an immunogenic composition for the treatment or

XX prevention of a Chlamydia trachomatis infection. The immunogenic

XX composition contains a Chlamydia trachomatis protein or its encoding

XX nucleic acid. The immunogenic composition of the invention is useful for

XX treating or preventing chlamydial infection. The present amino acid

XX sequence represents a Chlamydia trachomatis protein of the invention.

Best Local Similarity 97.0%; Pred. No. 0;  
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;  
QY 1 MQTSPHKFPLSMILAYSCCSLNGGYYAAEIMVPOGIYDGETLTVSFPYTVIGDPSGTTVF 60  
DB 1 MQTSPHKFPLSMILAYSCCSLNGGYYAAEIMVPOGIYDGETLTVSFPYTVIGDPSGTTVF 60  
QY 61 SAGELTLKNLONSIAALPLSCFGLNLLGSFTVLGRHSLTFENIRTSNGAALSAAADGL 120  
DB 61 SAGELTLKNLONSIAALPLSCFGLNLLGSFTVLGRHSLTFENIRTSNGAALSAAADGL 120  
QY 121 FTIEGFKELSFNCNLSLAVLPAATTNKGSTPTTSTPSNGTIYSKTDLLILNNEKFSF 180  
DB 121 FTIEGFKELSFNCNLSLAVLPAATTNKGSTPTTSTPSNGTIYSKTDLLILNNEKFSF 180  
QY 181 YSNLYSGGGGDAIDAKSLTVQGISKLCVFOENTAOADGACQVVTFSAMANEAPAFVAN 240  
DB 181 YSNLYSGGGGDAIDAKSLTVQGISKLCVFOENTAOADGACQVVTFSAMANEAPAFVAN 240  
QY 241 VAGVGGGIAAVQDQCGQVSSSTSTEDPVVSFSRNTAVEFDGNVARVGGIYSYGNVAPL 300  
DB 241 VAGVGGGIAAVQDQCGQVSSSTSTEDPVVSFSRNTAVEFDGNVARVGGIYSYGNVAPL 300  
QY 301 NNGKTLFLNNVASPVYIAAKOPTSCQASNTSNYGDGGAIFCKNGAQ-AGSNNSGSVSFD 359  
DB 301 NNGKTLFLNNVASPVYIAAKOPTSCQASNTSNYGDGGAIFCKNGAQ-AGSNNSGSVSFD 360  
QY 360 GEGVYVFFSSNVAAGKGGAIYAKKLSVANGCPVQFIENIANDGGAIVLGSBGLSADYG 419  
DB 361 GEGVYVFFSSNVAAGKGGAIYAKKLSVANGCPVQFIENIANDGGAIVLGSBGLSADYG 420  
QY 420 DIIIPGNLKRITAKENAADVNGVTVSSQAISSMGSGGKITTLRAKAGHQLLNDPIEMANG 479  
DB 421 DIIIPGNLKRITAKENAADVNGVTVSSQAISSMGSGGKITTLRAKAGHQLLNDPIEMANG 480  
QY 480 NQPAQSSBPLKINDGEGYTDIVFANGNSTLYQNTVEQGRIVLREKAKLSVNSLSQTGG 539  
DB 481 NQPAQSSBPLKINDGEGYTDIVFANGNSTLYQNTVEQGRIVLREKAKLSVNSLSQTGG 540  
QY 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNVNTPTPPAODSHP 599  
DB 541 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNVNTPTPPAODSHP 600  
QY 600 AVIGSTTAGSVTISGPIPFEDDDTDYDRYDWLGSNQKINVLKQLGKTGPANAPSDTL 659  
DB 601 AVIGSTTAGSVTISGPIPFEDDDTDYDRYDWLGSNQKIDVLKQLGQTPSANAPSDTL 660  
QY 660 GNEMPKYGGQSWKLAWDPTNANGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDI 719  
DB 661 GNEMPKYGGQSWKLAWDPTNANGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDI 720  
QY 720 RSHSAIQASVDRGSRVCRGLWVSGVSNPFYHDDALGQGYRISGGYSIGANSYFGSSMF 779  
DB 721 RSHSAIQASVDRGSRVCRGLWVSGVSNPFYHDDALGQGYRISGGYSIGANSYFGSSMF 780  
QY 780 GLAFTEVFGSRKDYVVCVCRSNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMKT 839  
DB 781 GLAFTEVFGSRKDYVVCVCRSNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMKT 840  
QY 840 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFSADHESFTEGD 899  
DB 841 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFSADHESFTEGD 900  
QY 900 QARAFKSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 959  
DB 901 QARAFRSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 960  
QY 960 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDSRGYLSAGSKVRFP 1012  
DB 961 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDSRGYLSAGSKVRFP 1013

RESULT 15

Query Match 97.4%; Score 5131.5; DB 9; Length 1013;

```
AEAL19080
ID AEAL19080 standard; protein; 1013 AA.
XX
AC AEAL19080;
XX
28-JUL-2005 (first entry)
XX
Chlamydia trachomatis protein - SEQ ID 83.
XX
Chlamydia trachomatis infection; antibacterial; vaccine.
XX
Chlamydia trachomatis.
XX
US2005106162-A1.
XX
19-MAY-2005.
XX
16-DEC-2004; 2004US-00498327.
XX
12-DEC-2001; 2001GB-00029732.
XX
06-AUG-2002; 2002EP-00182233.
XX
14-AUG-2002; 2001EP-00218924.
XX
12-DEC-2002; 2002WO-IB005761.
XX
(Grandi G.
PA (Ratti G.
XX
Grandi G, Ratti G;
XX
N-PSDB; AEAL19081.
XX
WPI; 2005-354777/36.
XX
DR DR
DR N-PSDB; AEAL19081.
XX
New immunogenic composition comprising Chlamydia trachomatis protein or
PT its encoding nucleic acid, useful for immunizing against, or treating or
PT preventing chlamydial infection.
XX
Claim 13; SEQ ID NO 83; 90pp; English.
XX
The invention comprises an immunogenic composition for the treatment or
CC prevention of a Chlamydia trachomatis infection. The immunogenic
CC composition contains a Chlamydia trachomatis protein or its encoding
CC nucleic acid. The immunogenic composition of the invention is useful
CC for treating or preventing chlamydial infection. The present amino acid
CC sequence represents a Chlamydia trachomatis protein of the invention.
CC NOTE: The present sequence is not shown in the specification, but was
CC obtained from the USPTO website -
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX
SQ Sequence 1013 AA;
XX
Query Match 97.3%; Score 5125.5; DB 9; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 21; Mismatches 11; Indels 1; Gaps 1;
QY 1 MQTSFHKFLLMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPPTVIGDPSGTTVF 60
DB 1 MQTSFHKFLLMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPPTVIGDPSGTTVF 60
QY 61 SAGELTLNKLNSIAALPLSCFNLGSLTFTVLRGHSITFTENRTSTNGAALSADGL 120
DB 61 SAGELTLNKLNSIAALPLSCFNLGSLTFTVLRGHSITFTENRTSTNGAALSADGL 120
QY 121 FTIEGFKELSPNCNLSLAVLPAATNKGSGTFTTSTPSNGTIYSKTDLLLNNKPSF 180
DB 121 FTIEGFKELSPNCNLSLAVLPAATNKGSGTFTTSTPSNGTIYSKTDLLLNNKPSF 180
QY 181 YSNLVSGDGGDAIDAKSLTVQIGSKLCVQENTAQADGGACQVTFSPSANEAPTAFAV 240
DB 181 YSNLVSGDGGDAIDAKSLTVQIGSKLCVQENTAQADGGACQVTFSPSANEAPTAFAV 240
QY 241 VAGVGGGGIAAVQDQCGQVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
DB 241 VAGVGGGGIAAVQDQCGQVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
```

```
QY 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYDGGAI FCKNGAQ-AGSNNSGSVSPD 359
DB 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSDNYDGGAI FCKNGAQAGSNNSGSVSPD 360
QY 360 GEGVVFSSNVAAGKGAIIYAKKLSVANCPGVQPLRNANDGGAIYIGESGELSADYG 419
DB 361 GEGVVFSSNVAAGKGAIIYAKKLSVANCPGVQPLRNANDGGAIYIGESGELSADYG 420
QY 420 DIIFDGNLKTAKENAADVNGVTVSSQAI SMGSGKITTIRAKAGHQLFNDPIEMANGN 479
DB 421 DIIFDGNLKTAKENAADVNGVTVSSQAI SMGSGKITTIRAKAGHQLFNDPIEMANGN 480
QY 480 NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTIQGRIVLREKAKLSVNSLSQTGG 539
DB 481 NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTIQGRIVLREKAKLSVNSLSQTGG 540
QY 540 SLYMEAGSTWDFVTPOPPQPPAANQLITLSNHLHLSLSSLLANNAVTPNPAPQDSHP 599
DB 541 SLYMEAGSTWDFVTPOPPQPPAANQLITLSNHLHLSLSSLLANNAVTPNPAPQDSHP 600
QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKINVLKIQGTGPPANAPSDLTL 659
DB 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKINVLKIQGTGPPANAPSDLTL 660
QY 660 GNEMPKYGYQGSWKLAWDPTANNPPTLKATWTKGTGNPQPRVAVSLVNSLWGSILDI 719
DB 661 GNEMPKYGYQGSWKLAWDPTANNPPTLKATWTKGTGNPQPRVAVSLVNSLWGSILDI 720
QY 720 RSAHSAIQASVDGRSYCRGLWVSGVSNPPYHDDRALGQGYRISGGYSLGANSYFGSSMF 779
DB 721 RSAHSAIQASVDGRSYCRGLWVSGVSNPPYHDDRALGQGYRISGGYSLGANSYFGSSMF 780
QY 780 GLAFTEVFGSRKDYVVCRSNHHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 839
DB 781 GLAFTEVFGSRKDYVVCRSNHHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 840
QY 840 SYTPAESDVRWNNCLAGBIAGLPIVITPSKLYLNLPPFVQAEPSYADHESFTREGD 899
DB 841 SYTPAESDVRWNNCLAGBIAGLPIVITPSKLYLNLPPFVQAEPSYADHESFTREGD 900
QY 900 QARAFKSGHLLNLSVPVGVKFDRCSSTHPNKYSFMAAII CDAYRTISGTETLLSHOETW 959
DB 901 QARAFKSGHLLNLSVPVGVKFDRCSSTHPNKYSFMAAII CDAYRTISGTETLLSHOETW 960
QY 960 TTDAFHLARHGVI VVGRGSMYASLTNSIEVYGHGRYERDASRGYGLSAGSVRF 1012
DB 961 TTDAFHLARHGVI VVGRGSMYASLTNSIEVYGHGRYERDASRGYGLSAGSVRF 1013
```

RESULT 16  
AEAL19006  
ID AEAL19006 standard; protein; 1013 AA.

XX AC AEAL19006;  
XX XX  
XX DT 28-JUL-2005 (first entry)  
XX XX  
XX DE Chlamydia trachomatis protein - SEQ ID 9.  
XX KW Chlamydia trachomatis infection; antibacterial; vaccine.  
XX OS Chlamydia trachomatis.  
XX XX  
XX FN US2005106162-A1.  
XX XX  
XX PD 19-MAY-2005.  
XX XX  
XX PF 16-DEC-2004; 2004US-00498327.  
XX PR 12-DEC-2001; 2001GB-00029732.  
XX PR 06-AUG-2002; 2002EP-00182233.  
XX PR 14-AUG-2002; 2001EP-00218924.



CC NOTE: The present sequence is not shown in the specification, but was  
CC obtained from the USPTO website -  
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.  
XX  
SQ Sequence 1013 AA;

Query Match	97.3%;	Score 5123.5;	DB 9;	Length 1013;
Best Local Similarity	96.6%;	Pred. No. 0;		
Matches	979;	Conservative	22;	Mismatches 11; Indels 1; Gaps 1;
Qy	1	MQTSFHKFFLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF	60	
Db	1	MQTSFHKFFLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF	60	
Qy	61	SAGELTLKNLNSIAALPLSCFNLGSLTFVLGRGHSITFTENIRTSNKAALSADGL	120	
Db	61	SAGELTLKNLNSIAALPLSCFNLGSLTFVLGRGHSITFTENIRTSNKAALSADGL	120	
Qy	121	FTIEGPKELSPNCNLSLAVLPAATNKGSTPTTTSTPSNGTIYSKTDLLILNNEKFSF	180	
Db	121	FTIEGPKELSPNCNLSLAVLPAATNKGSTPTTTSTPSNGTIYSKTDLLILNNEKFSF	180	
Qy	181	YSNLVSGDGAIDAKSLTVQGISKLCVQENTAOADGGACOVVTSFSAMANEAPAFVAN	240	
Db	181	YSNLVSGDGAIDAKSLTVQGISKLCVQENTAOADGGACOVVTSFSAMANEAPAFVAN	240	
Qy	241	VAGVGGGIAAVQDQGGVSGSTSTEDPVVPSRNTAVEFDGNVARVGGGIYSYGNVAF	300	
Db	241	VAGVGGGIAAVQDQGGVSGSTSTEDPVVPSRNTAVEFDGNVARVGGGIYSYGNVAF	300	
Qy	301	NGKTLFLNNVASPVYIAAKQTSOAGNTSNNGDGAIFCKNGAQ-AGSNNSGSVDF	359	
Db	301	NGKTLFLNNVASPVYIAAKQTSOAGNTSNNGDGAIFCKNGAQ-AGSNNSGSVDF	359	
Qy	360	GEVGVFFSNVAAGGAIYAKKLSVANGCPVQFLRNTANDGAIYLGESGELSADYG	419	
Db	361	GEVGVFFSNVAAGGAIYAKKLSVANGCPVQFLRNTANDGAIYLGESGELSADYG	420	
Qy	420	DIIPGNLKRKTAKENAADVGVTVSSQASISGSGGKITTLRAKAGHOILFNDPIEMANG	479	
Db	421	DIIPGNLKRKTAKENAADVGVTVSSQASISGSGGKITTLRAKAGHOILFNDPIEMANG	480	
Qy	480	NPAQSSKLLKINDEGVTGDIVFANGSSTLYQNTVIEQGRIVLBEKAKLSVNSLSQGG	539	
Db	481	NPAQSSKLLKINDEGVTGDIVFANGSSTLYQNTVIEQGRIVLBEKAKLSVNSLSQGG	540	
Qy	540	SLYMEAGTMDPVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP	599	
Db	541	SLYMEAGTMDPVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP	600	
Qy	600	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNKNVILKQLGTPKPPANAPSDTL	659	
Db	601	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNKNVILKQLGTPKPPANAPSDTL	660	
Qy	660	GNEMPKYQGSQKWLAMPNTANNPPTLKATWTGTGYNPGRVAVLSPNLSGSLDI	719	
Db	661	GNEMPKYQGSQKWLAMPNTANNPPTLKATWTGTGYNPGRVAVLSPNLSGSLDI	720	
Qy	720	RSASIAQASVDGRSVCRLWGVSNFFYHDDRDLGQYRISGGYSLGANSYFGSSMF	779	
Db	721	RSASIAQASVDGRSVCRLWGVSNFFYHDDRDLGQYRISGGYSLGANSYFGSSMF	780	
Qy	780	GLAFTEVFGSKDYVVCNRHACIGSVYLSLQALCGSYLFGDAFIRASYGFGNQHKMT	839	
Db	781	GLAFTEVFGSKDYVVCNRHACIGSVYLSLQALCGSYLFGDAFIRASYGFGNQHKMT	840	
Qy	840	SYTFAESDVWDDNNCLAGEIGAGLPIVITPSKLYNELRFPVQABFSAADHESFTEGD	899	
Db	841	SYTFAESDVWDDNNCLAGEIGAGLPIVITPSKLYNELRFPVQABFSAADHESFTEGD	900	
Qy	900	QARAFKSGHLNLSVPVGVKFDRCSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	959	
Db	901	QARAFKSGHLNLSVPVGVKFDRCSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	960	

Qy	960	TTDAFHLARHGVVVRGSMYASLTNSNIEVYHGRYEYRDASRGYLSAGSRVRF	1012	
Db	961	TTDAFHLARHGVVVRGSMYASLTNSNIEVYHGRYEYRDASRGYLSAGSRVRF	1013	

RESULT 18  
ID AEA19010 standard; protein; 1013 AA.  
XX AC AEA19010;  
XX 28-JUL-2005 (first entry)  
XX Chlamydia trachomatis protein - SEQ ID 13.  
XX chlamydia trachomatis infection; antibacterial; vaccine.  
XX Chlamydia trachomatis.  
XX OS US2005106162-A1.  
XX 19-MAY-2005.  
XX 16-DEC-2004; 2004US-00498327.  
XX 12-DEC-2001; 2001GB-00029732.  
XX 06-AUG-2002; 2002EP-00182233.  
XX 14-AUG-2002; 2001EP-00218924.  
XX 12-DEC-2002; 2002WO-1B005761.  
XX (GRAN/) GRANDI G.  
XX (RATT/) RATTI G.  
XX Grandi G, Ratti G;  
XX WPI; 2005-354777/36.  
XX N-PSDB; AEA19011.  
XX New immunogenic composition comprising Chlamydia trachomatis protein or its encoding nucleic acid, useful for immunizing against, or treating or preventing chlamydial infection.  
XX Claim 13; SEQ ID NO 13; 90pp; English.  
XX The invention comprises an immunogenic composition for the treatment or prevention of a Chlamydia trachomatis infection. The immunogenic composition contains a Chlamydia trachomatis protein or its encoding nucleic acid. The immunogenic composition of the invention is useful for treating or preventing chlamydial infection. The present amino acid sequence represents a Chlamydia trachomatis protein of the invention.  
XX NOTE: The present sequence is not shown in the specification, but was obtained from the USPTO website -  
XX seqdata.uspto.gov/sequence.html?DocID=20050106162.

Qy	1	MQTSFHKFFLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF	60	
Db	1	MQTSFHKFFLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF	60	
Qy	61	SAGELTLKNLNSIAALPLSCFNLGSLTFVLGRGHSITFTENIRTSNKAALSADGL	120	
Db	61	SAGELTLKNLNSIAALPLSCFNLGSLTFVLGRGHSITFTENIRTSNKAALSADGL	120	
Qy	121	FTIEGPKELSPNCNLSLAVLPAATNKGSTPTTTSTPSNGTIYSKTDLLILNNEKFSF	180	
Db	121	FTIEGPKELSPNCNLSLAVLPAATNKGSTPTTTSTPSNGTIYSKTDLLILNNEKFSF	180	

181	YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTFSFMANEAPIAFVAN	240
181	YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVVTFSFMANEAPIAFIAN	240
241	VAGVRGGGIAAVDQGGQVSSSTSTEDPVVFSRNTAVFDFGNVAVRGVGGIYSYGNVAF	300
241	VAGVRGGGIAAVDQGGQVSSSTSTEDPVVFSRNTAVFDFGNVAVRGVGGIYSYGNVAF	300
301	NRGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYDGGAIKCKNGAQ-AGSNNSGSVDF	359
301	NRGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYDGGAIKCKNGAQAGSNNSGSVDF	360
360	GGVVPFSSNVAAGKGAIIYAKKLSVANGCPVQFLRNINANDGGAIIYLGESGELSADYG	419
361	GGVVPFSSNVAAGKGAIIYAKKLSVANGCPVQFLRNINANDGGAIIYLGESGELSADYG	420
420	DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLIFNDPIEMANG	479
421	DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLIFNDPIEMANG	480
480	NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG	539
481	NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG	540
540	SLYMEAGSTWDFVTPPOPPQPPAANOLITLSNHLSSLANNVNTPTNPPAQDSHP	599
541	SLYMEAGSTWDFVTPPOPPQPPAANOLITLSNHLSSLANNVNTPTNPPAQDSHP	600
600	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWMGSKINVLKQLGTPKANAPSDTLT	659
601	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWMGSKINVLKQLGTPKANAPSDTLT	660
660	GNEMPKYGVGSKLAWDPNTANGPYTLKATWTKGYNPGPERSVSLVNSLWGSILDI	719
661	GNEMPKYGVGSKLAWDPNTANGPYTLKATWTKGYNPGPERSVSLVNSLWGSILDI	720
720	RSASHAIQASVDGRSVCRLGWGVSNNFFVHDDALDQGVYIYSGGYSGLANSYFGSSMF	779
721	RSASHAIQASVDGRSVCRLGWGVSNNFFVHDDALDQGVYIYSGGYSGLANSYFGSSMF	780
780	GLAFTEVFRGSKDYVCRSNHACIGSVYLSLSTQALCGSLYFGDAFLRASYGFGNQHMT	839
781	GLAFTEVFRGSKDYVCRSNHACIGSVYLSLSTQALCGSLYFGDAFLRASYGFGNQHMT	840
840	SYTFAESDVRDNNCLAGBIGAGLPITVTPSKLYLNELRPFVQAFPSYADHESFTEEGD	899
841	SYTFAESDVRDNNCLAGBIGAGLPITVTPSKLYLNELRPFVQAFPSYADHESFTEEGD	900
900	QARAFKSHLLNLSPVGVKFDRCSSVTHPNKYSPMAAYICDAYRTISGTTLLSHQETW	959
901	QARAFKSHLLNLSPVGVKFDRCSSVTHPNKYSPMAAYICDAYRTISGTTLLSHQETW	960
960	TTDAFLHARHGVVRGSMYASLTSNIEVYGHGRYEDRSDRGYGLSAGSRVRF	1012
961	TTDAFLHARHGVVRGSMYASLTSNIEVYGHGRYEDRSDRGYGLSAGSRVRF	1013

RESULT 19

EA19000

D AEA19000 standard; protein; 1013 AA.

X AEA19000;

C AEA19000;

X 28-JUL-2005 (first entry)

X Chlamydia trachomatis protein - SEQ ID 3.

X Chlamydia trachomatis infection; antibacterial; vaccine.

X Chlamydia trachomatis.

X Chlamydia trachomatis.

X US2005106162-A1.

Db 481 NPAQSSBPLKINDGEGYTDIVFANGNSTLYQNVTTIEQGRIVLRKAKLSVNSLSQTGG 540  
 Qy 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599  
 Db 541 SLYMEAGSTLDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 600  
 Qy 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKINVLKQLGTPKPNAPSDDLTL 659  
 Db 601 AIIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKIDVLKQLGTPSANAPSDDLTL 660  
 Qy 660 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTKTGYNPQPERVASLVPNSLWGSILDI 719  
 Db 661 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTKTGYNPQPERVASLVPNSLWGSILDI 720  
 Qy 720 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRYISGGYSLGANSYFGSSMF 779  
 Db 721 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRYISGGYSLGANSYFGSSMF 780  
 Qy 780 GLAFTEVFGRSKDYVVCBSNHHACIGSVYLSLQALCGSYLFGDAFIRASVYFGNQHMKT 839  
 Db 781 GLAFTEVFGRSKDYVVCBSNHHACIGSVYLSLQALCGSYLFGDAFLRASVYFGNQHMKT 840  
 Qy 840 SYTFAEEDVRWNNCLAGEIGAGLPIVITPSSKLYINELRPVQAEPSYADHESFTREGD 899  
 Db 841 SYTFAEEDVRWNNCLVGEIGVGLPIVITPSSKLYINELRPVQAEPSYADHESFTREGD 900  
 Qy 900 QARAFKSHLLNLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959  
 Db 901 QARAFKSHLLNLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960  
 Qy 960 TTDAPHLARHGVIYVGRSMYASLTSNIYVGHGREYRDASRGYGLSAGSVRF 1012  
 Db 961 TTDAPHLARHGVIYVGRSMYASLTSNIYVGHGREYRDTSRGYGLSAGSKVRF 1013

## RESULT 20

AEA18998  
 ID AEA18998 standard; protein; 1013 AA.  
 AC AEA18998;  
 DT 28-JUL-2005 (first entry)  
 XX Chlamydia trachomatis protein - SEQ ID 1.  
 XX Chlamydia trachomatis infection; antibacterial; vaccine.  
 XX Chlamydia trachomatis.  
 XX US2005106162-A1.  
 XX 19-MAY-2005.  
 XX 16-DEC-2004; 2004US-00498327.  
 XX 12-DEC-2001; 2001GB-00029732.  
 PR 06-AUG-2002; 2002EP-00182233.  
 PR 14-AUG-2002; 2001EP-00218924.  
 PR 12-DEC-2002; 2002WO-IB005761.  
 XX (GRAN/) GRANDI G.  
 PA (RATT/) RATTI G.  
 XX Grandi G, Ratti G;  
 XX MPI; 2005-354777/36.  
 DR N-P8DB; AEA18999.  
 XX New immunogenic composition comprising Chlamydia trachomatis protein or  
 PT its encoding nucleic acid, useful for immunizing against, or treating or  
 PT preventing chlamydial infection.  
 XX Claim 13; SEQ ID NO 1; 90pp; English.

XX The invention comprises an immunogenic composition for the treatment or  
 CC prevention of a Chlamydia trachomatis infection. The immunogenic  
 CC composition contains a Chlamydia trachomatis protein or its encoding  
 CC nucleic acid. The immunogenic composition of the invention is useful for  
 CC treating or preventing chlamydial infection. The present amino acid  
 CC sequence represents a Chlamydia trachomatis protein of the invention.  
 CC NOTE: The present sequence is not shown in the specification, but was  
 CC obtained from the USPTO website -  
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.

XX Sequence 1013 AA;

Query Match 97.2%; Score 5121.5; DB 9; Length 1013;  
 Best Local Similarity 96.5%; Pred. No. 0;  
 Matches 978; Conservative 23; Mismatches 11; Indels 1; Gaps 1;

Qy 1 MQTSFHKPFLSMILAYSCSLNGGYYAARIMVPGIYDGETLTVSPVTVIGDPSGTTVF 60  
 Db 1 MQTSFHKPFLSMILAYSCSLNGGYYAARIMVPGIYDGETLTVSPVTVIGDPSGTTVF 60  
 Qy 61 SAGELTLKNDNSIAALPLSCFGLGSLFTVLRGHSHTFENIRTSINGAALSAAADGL 120  
 Db 61 SAGELTLKNDNSIAALPLSCFGLGSLFTVLRGHSHTFENIRTSINGAALSAAADGL 120  
 Qy 121 FTIRGPKELSFNSCNLSLAVLPAATTNKGSTPTTSTPSTNGTIYSKTDLLLNNKPSF 180  
 Db 121 FTIRGPKELSFNSCNLSLAVLPAATTNKGSTPTTSTPSTNGTIYSKTDLLLNNKPSF 180  
 Qy 181 YSNLVSGDGGAIKAKSLTVQGISKLCYPOENTAQDGGACOVVTSFSAMANEAPAVAN 240  
 Db 181 YSNLVSGDGGAIKAKSLTVQGISKLCYPOENTAQDGGACOVVTSFSAMANEAPAVAN 240  
 Qy 241 VAGVGGGIAAVDQGGVSSSTSTEDPVVSFSNTAVFEDGNVAVGGGIYSYGNVAF 300  
 Db 241 VAGVGGGIAAVDQGGVSSSTSTEDPVVSFSNTAVFEDGNVAVGGGIYSYGNVAF 300  
 Qy 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYDGGAIKAKGAGQILFNDPIEMANG 479  
 Db 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYDGGAIKAKGAGQILFNDPIEMANG 479  
 Qy 360 GEGVVPFSSNVAAGKGAIVAKKLSVANCGVPQPLRNINDGGAIYIGESGELSADYG 419  
 Db 360 GEGVVPFSSNVAAGKGAIVAKKLSVANCGVPQPLRNINDGGAIYIGESGELSADYG 419  
 Qy 420 DIIFDGNLKRKTAKENAADVNGVTVSSQAIKSGSGGKITTLRAKAGHQLFNDPIEMANG 479  
 Db 420 DIIFDGNLKRKTAKENAADVNGVTVSSQAIKSGSGGKITTLRAKAGHQLFNDPIEMANG 479  
 Qy 480 NQPAQSSBPLKINDGEGYTDIVFANGNSTLYQNVTTIEQGRIVLRKAKLSVNSLSQTGG 539  
 Db 480 NQPAQSSBPLKINDGEGYTDIVFANGNSTLYQNVTTIEQGRIVLRKAKLSVNSLSQTGG 540  
 Qy 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599  
 Db 541 SLYMEAGSTLDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 600  
 Qy 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKINVLKQLGTPKPNAPSDDLTL 659  
 Db 601 AIIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKIDVLKQLGTPSANAPSDDLTL 660  
 Qy 660 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTKTGYNPQPERVASLVPNSLWGSILDI 719  
 Db 661 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTKTGYNPQPERVASLVPNSLWGSILDI 720  
 Qy 720 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRYISGGYSLGANSYFGSSMF 779  
 Db 721 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRYISGGYSLGANSYFGSSMF 780  
 Qy 780 GLAFTEVFGRSKDYVVCBSNHHACIGSVYLSLQALCGSYLFGDAFIRASVYFGNQHMKT 839  
 Db 781 GLAFTEVFGRSKDYVVCBSNHHACIGSVYLSLQALCGSYLFGDAFLRASVYFGNQHMKT 840



QY 840 SYTFAESDVRDNNCLAGEIGAGLPVITPSKLYLNEIRPFVQAEFSYADHESFTEGD 899  
DB 841 SYTFAESDVRDNNCLVGEIGVGLPVIITPSKLYLNEIRPFVQAEFSYADHESFTEGD 900  
QY 900 QARAFKSGHLLNLSPVGVKDFRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959  
DB 901 QARAFKSGHLLNLSPVGVKDFRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960  
QY 960 TTDAFHLARHGVIIVRGSVMYASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF 1012  
DB 961 TTDAFHLARHGVIIVRGSVMYASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF 1013

RESULT 21  
ID AEA19014 standard; protein; 1013 AA.  
AC AEA19014;  
XX 28-JUL-2005 (first entry)  
DE Chlamydia trachomatis protein - SEQ ID 17.  
XX chlamydia trachomatis infection; antibacterial; vaccine.  
XX Chlamydia trachomatis.  
XX US2005106162-A1.  
XX 19-MAY-2005.  
XX 16-DEC-2004; 2004US-00498327.  
XX 12-DEC-2001; 2001GB-00029732.  
PR 06-AUG-2002; 2002EP-00182233.  
PR 14-AUG-2002; 2001EP-00218924.  
PR 12-DEC-2002; 2002WO-1B005761.  
XX (GRAN/) GRANDI G.  
PA (RATT/) RATTI G.  
XX Grandi G, Ratti G;  
XX WPI; 2005-354777/36.  
DR N-PSDB; AEA19015.  
XX New immunogenic composition comprising Chlamydia trachomatis protein or its encoding nucleic acid, useful for immunizing against, or treating or preventing chlamydial infection.  
XX Claim 13; SEQ ID NO 17; 90pp; English.  
XX The invention comprises an immunogenic composition for the treatment or prevention of a Chlamydia trachomatis infection. The immunogenic composition contains a Chlamydia trachomatis protein or its encoding nucleic acid. The immunogenic composition of the invention is useful for treating or preventing chlamydial infection. The present amino acid sequence represents a Chlamydia trachomatis protein of the invention.  
XX NOTE: The present sequence is not shown in the specification, but was obtained from the USPTO website -  
XX seqdata.uspto.gov/sequence.html?docID=20050106162.  
XX Sequence 1013 AA;

Query Match 97.1%; Score 5112.5; DB 9; Length 1013;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 977; Conservative 23; Mismatches 12; Indels 1; Gaps 1;  
QY 1 MQTSFHKFFLLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTTVF 60  
DB 1 MQTSFHKFFLLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTTVF 60  
QY 61 SAGELTLKNDLSIALPLSCFGNLGSFTVLGRGHSLTFTENRTSTNGAALSANDGL 120

DB 61 SAGELTLKNDLSIALPLSCFGNLGSFTVLGRGHSLTFTENRTSTNGAALSANDGL 120  
QY 121 FTIEGFKELSFSCNCSLLAVLPAATTNKGSTPTTSTPNSNGTIYSKTDLLLLNKEFSF 180  
DB 121 FTIEGFKELSFSCNCSLLAVLPAATTNKGSTPTTSTPNSNGTIYSKTDLLLLNKEFSF 180  
QY 181 YSNLVSDGGDAIDAKSLTVQGISKLCVQFQNTAQADGGACQVVTFSFMANEAPIAFVAN 240  
DB 181 YSNLVSDGGDAIDAKSLTVQGISKLCVQFQNTAQADGGACQVVTFSFMANEAPIAFVAN 240  
QY 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVFSFNTAVEFDGNVAVCGGIIYSYGNVAPL 300  
DB 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVFSFNTAVEFDGNVAVCGGIIYSYGNVAPL 300  
QY 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNTYDGGAIKCKGGAQ-AGSNNSGSVSPD 359  
DB 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNTYDGGAIKCKGGAQAGSNNSGSVSPD 360  
QY 360 GEGVVFSSNVAAGKGGAIYAKLSVANCGPVQFLNLTANDGGAIYLGESGELSADYG 419  
DB 361 GEGVVFSSNVAAGKGGAIYAKLSVANCGPVQFLNLTANDGGAIYLGESGELSADYG 420  
QY 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGKITTLRAKAGHQLFNDPIEMANGN 479  
DB 421 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGKITTLRAKAGHQLFNDPIEMANGN 480  
QY 480 NPAQSSKLLAKINDGEGYTDIVFANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSQTCG 539  
DB 481 NPAQSSKLLAKINDGEGYTDIVFANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSQTCG 540  
QY 540 SLYMEAGSTWDFVTPQQPPQPAANQLITLSNLHLSSLANNATVTPPTPPAQDSHP 599  
DB 541 SLYMEAGSTWDFVTPQQPPQPAANQLITLSNLHLSSLANNATVTPPTPPAQDSHP 600  
QY 600 AVIGSTTAGSVTISGPIFPEDLDDTAYDRYDWLGSNKNIVLKLQICYPANASDLTL 659  
DB 601 AIIGSTTAGSVTISGPIFPEDLDDTAYDRYDWLGSNKNIVLKLQICYPANASDLTL 660  
QY 660 GNEMPKYGYQGSWKLAWDPNTANNPYTLKATWTKTGYNPGPVERVASLVPNSLWSILDI 719  
DB 661 GNEMPKYGYQGSWKLAWDPNTANNPYTLKATWTKTGYNPGPVERVASLVPNSLWSILDI 720  
QY 720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDALGQGYRISGYSIGANSYFGSSMF 779  
DB 721 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDALGQGYRISGYSIGANSYFGSSMF 780  
QY 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKT 839  
DB 781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKT 840  
QY 840 SYTFAESDVRDNNCLAGEIGAGLPVITPSKLYLNEIRPFVQAEFSYADHESFTEGD 899  
DB 841 SYTFAESDVRDNNCLVGEIGVGLPVIITPSKLYLNEIRPFVQAEFSYADHESFTEGD 900  
QY 900 QARAFKSGHLLNLSPVGVKDFRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959  
DB 901 QARAFKSGHLLNLSPVGVKDFRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960  
QY 960 TTDAFHLARHGVIIVRGSVMYASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF 1012  
DB 961 TTDAFHLARHGVIIVRGSVMYASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF 1013

RESULT 22  
ID AAY16737 standard; protein; 1013 AA.  
XX AAY16737;  
XX 21-JUL-1999 (first entry)  
XX C. trachomatis B serovar HMW protein.

```

XX Chlamydia; high molecular weight protein; HMW protein; urethritis;
KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
KW salpingitis; tubal occlusion; infertility; cervical cancer;
KW arteriosclerosis; atherosclerosis.
XX Chlamydia trachomatis.
OS MO9917741-A1.
XX PN 15-APR-1999.
XX PD 01-OCT-1998; 98WO-US020737.
XX PF 02-OCT-1997; 97US-00942596.
XX PR (ANTE-) ANTEX BIOLOGICS INC.
XX PA Jackson JW, Pace JL;
XX PI WPI, 1999-287659/24.
XX DR New Chlamydia protein useful for treating conjunctivitis, urethritis and
XX PT cervical cancer.
XX PS Claim 4, Page 115-119; 14pp; English.
XX CC The invention relates to an isolated Chlamydia species high molecular
CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
CC be used for preventing, treating or ameliorating a disorder related to
CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
CC cancer, infertility, arteriosclerosis and atherosclerosis. The products
CC can also be used for detection and diagnosis. The present sequence
CC represents a C. trachomatis HMW protein
XX Sequence 1013 AA;

Query Match 97.0%; Score 5106.5; DB 2; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;

QY 1 MQTSPHKPFLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPYTVIGDPSGTTVF 60

QY 61 SAGELTLKLNLSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGLSNAADGL 120
DB 61 SAGELTLKLNLSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGLSNAADGL 120

QY 121 FTIEGFKELSPNCNLSLAVLPAATNKGSGTPTTTPSNGTIYSKTDLLLLANKEPSF 180
DB 121 FTIEGFKELSPNCNLSLAVLPAATNKGSGTPTTTPSNGTIYSKTDLLLLANKEPSF 180

QY 181 YSNLVSGDGGDAIDAKSLTVQGISKLCVFOENTAQADGGACQVTSFSAMANEAPAFVAN 240
DB 181 YSNLVSGDGGDAIDAKSLTVQGISKLCVFOENTAQADGGACQVTSFSAMANEAPAFVAN 240

QY 241 VAGVRGGGIAAVDQGGQVSSSTSTEDPVVFSRNTAVFPGNVARVCGGIYSYGNVAPL 300
DB 241 VAGVRGGGIAAVDQGGQVSSSTSTEDPVVFSRNTAVFPGNVARVCGGIYSYGNVAPL 300

QY 301 NNGKTLFLNNVASPVYIAAKQPTSCAGNTSNNGYDGGAIKCKNGAQ-AGSNNGSVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAKQPTSCAGNTSNNGYDGGAIKCKNGAQ-AGSNNGSVSFD 359

QY 360 GEGVVFFSNNVAAAGKGAIIYAKLSVANGCPVQFLRNITANDGGAIIYLGESGELSISADYG 419
DB 361 GEGVVFFSNNVAAAGKGAIIYAKLSVANGCPVQFLRNITANDGGAIIYLGESGELSISADYG 420

420 DIIFDGNLKRKTAKENAADVNGVTVSSQAIISMGSGGKITTILRAKAGHOILFNDPIEMANGN 479
421 DMIIFDGNLKRKTAKENAADVNGVTVSSQAIISMGSGGKITTILRAKAGHOILFNDPIEMANGN 480
480 NQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQSGG 539
481 NQPAQSSEPLKINDGEGYTGDIIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQSGG 540
540 SLYMEAGSTWDFVTPPQPPQPPAANQILITLNLHLSLSSLLANNVNTNPPNPPAQDSHP 599
541 SLYMEAGSTLDFVTPPQPPQPPAANQISITLNLHLSLSSLLANNVNTNPPNPPAQDSHP 600
600 AVIGSTTAGSVTISGPIFFEDDDTAYDRYDMLGSNOKINVLQLOLTKPPANAPSDTL 659
601 AVIGSTTAGSVTISGPIFFEDDDTAYDRYDMLGSNOKIDVLQLOLTKPPANAPSDTL 660
660 GNEMPKYGYQGSWKLAWDPTNANNPPTLKAATWTKTGYNFQPERVASLVPSNLWGSILDI 719
661 GNEMPKYGYQGSWKLAWDPTNANNPPTLKAATWTKTGYNFQPERVASLVPSNLWGSILDI 720
720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFTHDRDALCGGYRIISGGYSLGANSYFGSSMF 779
721 RSAHSAIQASVDGRSYCRGLWVSGVSNFFTHDRDALCGGYRIISGGYSLGANSYFGSSMF 780
780 GLAFTEVFGRSKDYVVCRSNHHACIGSVLSTQOALCGSVLFGDAFIRASVGCNQHMKT 839
781 GLAFTEVFGRSKDYVVCRSNHHACIGSVLSTQOALCGSVLFGDAFIRASVGCNQHMKT 840
840 SYTPAESDVRWNNCLAGIBIGLPIVITPSPKLYLNLRLPFFVQAEFSYADHESFTREGD 899
841 SYTPAESDVRWNNCLAGIBIGLPIVITPSPKLYLNLRLPFFVQAEFSYADHESFTREGD 900
900 QARAFKSGHLLNLSPVGVGFDRCSSTHKNKYSMAAYICDAYTISGTETILLSHOBTW 959
901 QARAFKSGHLLNLSPVGVGFDRCSSTHKNKYSMAAYICDAYTISGTETILLSHOBTW 960
960 TTDAPHLARHGVVVRGSMVYASLTNSIEVGHGRYERDASRGYGLSAGSRVRF 1012
961 TTDAPHLARHGVVVRGSMVYASLTNSIEVGHGRYERDASRGYGLSAGSRVRF 1013

RESULT 23
AAV16738
ID AAV16738 standard; protein; 1013 AA.
XX
AC AAV16738;
XX
DT 21-JUL-1999 (first entry)
XX
DE C. trachomatis F serovar HMW protein.
XX
KW Chlamydia; high molecular weight protein; HMW protein; urethritis;
KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
KW salpingitis; tubal occlusion; infertility; cervical cancer;
KW arteriosclerosis; atherosclerosis.
XX
OS Chlamydia trachomatis.
XX
PN MO9917741-A1.
XX
PD 15-APR-1999.
XX
PF 01-OCT-1998; 98WO-US020737.
XX
PR 02-OCT-1997; 97US-00942596.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson JW, Pace JL;
XX
DR WPI, 1999-287659/24.
XX

```

PT New Chlamydia protein useful for treating conjunctivitis, urethritis and  
PT cervical cancer.  
XX  
XX  
XX Claim 4; Page 119-123; 141pp; English.  
XX  
XX The invention relates to an isolated Chlamydia species high molecular  
CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as  
CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can  
CC be used for preventing, treating or ameliorating a disorder related to  
CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,  
CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,  
CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical  
CC cancer, infertility, arteriosclerosis and atherosclerosis. The products  
CC can also be used for detection and diagnosis. The present sequence  
CC represents a C. trachomatis HMW protein  
XX  
XX Sequence 1013 AA;  
SQ

Query Match 97.0%; Score 5106.5; DB 2; Length 1013;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

QY 1 MQTSPHKFPLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPGTTVF 60  
DB 1 MQTSPHKFPLSMILAYSCSLNGGYYAAEIMVPOGIYDGETLTVSPFYTVIGDPGTTVF 60  
QY 61 SAGELTLKNLNSIAALPLSCFNLGSLTVLGRGHSITFENIRTSNGAALSNAADGL 120  
DB 61 SAGELTLKNLNSIAALPLSCFNLGSLTVLGRGHSITFENIRTSNGAALSNAADGL 120  
QY 121 FTIEGFKELFNCNLSLAVLAATNKGSOPTTSTPSNCTIYKTDLLLNNEKFSF 180  
DB 121 FTIEGFKELFNCNLSLAVLAATNKGSOPTTSTPSNCTIYKTDLLLNNEKFSF 180  
QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGACQVTVSFSAMANEAPAFVAN 240  
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGACQVTVSFSAMANEAPAFVAN 240  
QY 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVFEDGNVAVRGVGGIYSGNVAPL 300  
DB 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVFEDGNVAVRGVGGIYSGNVAPL 300  
QY 301 NNGKTLFLNNVASPVYIAAQKTSQASNTNNGYDGGAI FCKNGAQ-AGSNNSGSVSD 359  
DB 301 NNGKTLFLNNVASPVYIAAQKTSQASNTNNGYDGGAI FCKNGAQAGSNNSGSVSD 360  
QY 360 GEGWVFPSSNVAAGKGAIVAKLSVANGCPVQFLRIANDGGAIVLGSSELSDADYG 419  
DB 361 GEGWVFPSSNVAAGKGAIVAKLSVANGCPVQFLRIANDGGAIVLGSSELSDADYG 420  
QY 420 DIIFDGNLXRTAKENAAVNGVTVSSQAISMGSGGKIITTLRAKAGHOILFNDPIEMANG 479  
DB 421 DIIFDGNLXRTAKENAAVNGVTVSSQAISMGSGGKIITTLRAKAGHOILFNDPIEMANG 480  
QY 480 NQPAQSKLLKINDGGYTGDIIVFANGSSTLYQNVTIEQRIIVLRKAKLSVNSLSOTGG 539  
DB 481 NQPAQSKLLKINDGGYTGDIIVFANGSSTLYQNVTIEQRIIVLRKAKLSVNSLSOTGG 540  
QY 540 SLYMEAGSTWDFVTPPOPPQPPAANOLITLSNLHLSLLANNVTPNTPPEADQSH 599  
DB 541 SLYMEAGSTWDFVTPPOPPQPPAANOLITLSNLHLSLLANNVTPNTPPEADQSH 600  
QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKNKLNLKQLGTPPPANAPSDLT 659  
DB 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKNKLNLKQLGTPPPANAPSDLT 660  
QY 660 GNEMPKYGYQGSWKLAWDPNTANNPGPYTLKATWTGTGYNPGERVASLVNSLWGLDI 719  
DB 661 GNEMPKYGYQGSWKLAWDPNTANNPGPYTLKATWTGTGYNPGERVASLVNSLWGLDI 720  
QY 720 RSAHSAIQASVDCRSYCRGLWVGVSNNFFVHDDALGQGYRIYSGGYSLCANSYFGSSMF 779  
DB 721 RSAHSAIQASVDCRSYCRGLWVGVSNNFFVHDDALGQGYRIYSGGYSLCANSYFGSSMF 780

QY 780 GLAPTEVGRSKDYVVCRRNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 839  
DB 781 GLAPTEVGRSKDYVVCRRNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 840  
QY 840 SYTPAESDVRWDNCLVGEIGVGLPIVITPSPKUYLNLRLPFPVQAEFSYADHESFTEGD 899  
DB 841 SYTPAESDVRWDNCLVGEIGVGLPIVITPSPKUYLNLRLPFPVQAEFSYADHESFTEGD 900  
QY 900 QARAFKSHLLNLSVPVGVKFDRCSSHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959  
DB 901 QARAFKSHLLNLSVPVGVKFDRCSSHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960  
QY 960 TTDAPFLARHGVVVRGSMYASLTNIEVYHGRYEDASRGYGLSAGSRVRF 1012  
DB 961 TTDAPFLARHGVVVRGSMYASLTNIEVYHGRYEDASRGYGLSAGSRVRF 1013

RESULT 24  
AAG83207  
ID AAG83207 standard; protein; 1006 AA.  
XX  
AC AAG83207;  
XX  
DT 05-SEP-2001 (first entry)  
XX  
DE Protein encoded by Chlamydia trachomatis pmpG gene.  
XX  
KW Chlamydia; vaccine; infection; fusion protein; antigen;  
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;  
KW acute respiratory tract infection; Cap1; CTS29; OMCB;  
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.  
XX  
OS Chlamydia trachomatis.  
XX  
PN WO200140474-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 04-DEC-2000; 2000WO-US032919.  
XX  
PR 03-DEC-1999; 99US-00454684.  
PR 19-APR-2000; 2000US-00556877.  
XX 20-JUN-2000; 2000US-00598419.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Probst P, Bhatia A, Skeiky YAM, Fling SP, Scholler J;  
XX  
DR WPI; 2001-374831/39.  
XX  
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic  
PT inflammatory disease, trachoma, acute respiratory tract infections,  
PT atherosclerosis and heart disease.  
XX  
PS Claim 2; Page 214-216; 295pp; English.  
XX  
CC The present sequence is provided in a specification relating to compounds  
CC and methods for the treatment and diagnosis of chlamydial infection. The  
CC compounds provided include polypeptides and fusion proteins comprising  
CC immunogenic portions of Chlamydia antigens and DNA sequences encoding  
CC such polypeptides. They are useful for vaccinating against chlamydial  
CC infection, which causes pelvic inflammatory disease, trachoma, acute  
CC respiratory tract infections, atherosclerosis and heart disease  
XX  
SQ Sequence 1006 AA;  
Query Match 96.6%; Score 5090; DB 4; Length 1006;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 979; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 31 MVPOGIYDGETLTVSPFYTVIGDPGTTVFSAGELTLKNLNSIAALPLSCFNLGSPFT 90  
DB 31 MVPOGIYDGETLTVSPFYTVIGDPGTTVFSAGELTLKNLNSIAALPLSCFNLGSPFT 90

Db 25 MIPOGIYDGETLTVSPPTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFGNLGSFT 84

Qy 91 VLGRHSLTPENIRTSNGAALSAAADGLFTIEGFKELSFNSCNLSLAVLPAATTNKG 150

Db 85 VLGRHSLTPENIRTSNGAALSAAADGLFTIEGFKELSFNSCNLSLAVLPAATTNKG 144

Qy 151 QTPTTSTPSNGTIYSKTDLLLNNEKFSFYNLSVGGDAIDAKSLTVQGISKLCVFOE 210

Db 145 QTPTTSTPSNGTIYSKTDLLLNNEKFSFYNLSVGGDAIDAKSLTVQGISKLCVFOE 204

Qy 211 NTAQADGACQVTSFSAANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 270

Db 205 NTAQADGACQVTSFSAANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 264

Qy 271 SFSRNTAVFPGNVARVGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 330

Db 265 SFSRNTAVFPGNVARVGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 324

Qy 331 SNNTYDGGGAI FCKNGAQAGSNNSGVSPDGRGVVFPFSSNVAAGKGAIYAKKLSVANC 390

Db 325 SNNTYDGGGAI FCKNGAQAGSNNSGVSPDGRGVVFPFSSNVAAGKGAIYAKKLSVANC 384

Qy 391 VOFLRNANDGGAIYLGESBELSLADYGDII FGNLKRKTAKENAADVNGVTVSSQAISM 450

Db 385 VOFLRNANDGGAIYLGESBELSLADYGDII FGNLKRKTAKENAADVNGVTVSSQAISM 444

Qy 451 GSGGKITTLRAKAGHQLIFNDPIEMANGNNOQAQSSKLLKINDGEYTGDIIVPANGSSTL 510

Db 445 GSGGKITTLRAKAGHQLIFNDPIEMANGNNOQAQSSKLLKINDGEYTGDIIVPANGSSTL 504

Qy 511 YQNTVIEGRIVLREKAKLSVNSISOTGGSILYMEAGSTWDPVTPPOQPPAANQLITLS 570

Db 505 YQNTVIEGRIVLREKAKLSVNSISOTGGSILYMEAGSTWDPVTPPOQPPAANQLITLS 564

Qy 571 NLHLSLSLLANNAVTPPTPPAQDSHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDRYD 630

Db 565 NLHLSLSLLANNAVTPPTPPAQDSHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDRYD 624

Qy 631 WLGNQKINVLKQLGTPKANASBDLTGLNEMPKYGYQGSWKLAWDPNTANNPYTLKA 690

Db 625 WLGNQKINVLKQLGTPKANASBDLTGLNEMPKYGYQGSWKLAWDPNTANNPYTLKA 684

Qy 691 TWTKTGNRGPVRSVSLVPSNLMSIILDIRSAHSAIQASVDGRSVCRLWVSGVSNPFYH 750

Db 685 TWTKTGNRGPVRSVSLVPSNLMSIILDIRSAHSAIQASVDGRSVCRLWVSGVSNPFYH 744

Qy 751 DRDALGQGYRIISGGYSLGANSYFGSSMFGLAFTFVGRSKDYVVCRSNHACIGSVYLS 810

Db 745 DRDALGQGYRIISGGYSLGANSYFGSSMFGLAFTFVGRSKDYVVCRSNHACIGSVYLS 804

Qy 811 TQOALCGSYLPDGAFTIRASYFGNQHMTSYTPAEDSDVRDNNCLAGEIGAGLPVITP 870

Db 805 TQOALCGSYLPDGAFTIRASYFGNQHMTSYTPAEDSDVRDNNCLAGEIGAGLPVITP 864

Qy 871 SKLYNELRPVQAEPSYADHESFTEBGDQARAFKSGHLLNLSVPVGVKPRDCSSTHPNK 930

Db 865 SKLYNELRPVQAEPSYADHESFTEBGDQARAFKSGHLLNLSVPVGVKPRDCSSTHPNK 924

Qy 931 YSFMVAAYICDAYRTISGTETTLTLLSHOETWTWTDAPHARHGVVVRGSMYASLTSNIEVYGH 990

Db 925 YSFMVAAYICDAYRTISGTETTLTLLSHOETWTWTDAPHARHGVVVRGSMYASLTSNIEVYGH 984

Qy 991 GRYEYRDASRGYGLSAGSRVRF 1012

Db 985 GRYEYRDASRGYGLSAGSKVRP 1006

RESULT 25  
 ABB94178  
 ID ABB94178 standard; protein; 1006 AA.  
 XX  
 AC ABB94178;  
 XX

DT 05-JUN-2002 (first entry)

XX Chlamydia protein sequence SEQ ID NO:190.

DE Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;

XX antibacterial; immunostimulant; immune response;

KW Chlamydia-specific T-cell response.

KW Chlamydia sp.

OS WO200208267-A2.

PN 31-JAN-2002.

XX 20-JUL-2001; 2001WO-US023121.

PP 23-JUL-2000; 2000US-00620412.

XX 20-APR-2001; 2001US-00841132.

PR (CORI-) CORIXA CORP.

XX Fling SP, Skeiky YAW, Probat P, Bhatia A;

XX MPI; 2002-179901/23.

XX Novel compositions comprising Chlamydia Cap1 protein and its use in the treatment of Chlamydia infection.

XX Example 1; Page 245-247; 537pp; English.

PS The present invention describes compositions comprising a Chlamydia Cap1 protein and methods for the diagnosis and therapy of Chlamydia infection.

CC Chlamydia DNA and protein sequences from the present invention can have antibacterial and immunostimulant activities, and can be used in vaccines. Compounds from the present invention can be used for eliciting an immune response, specifically stimulating a Chlamydia-specific T-cell response or inhibiting the development of a Chlamydia infection in an animal. Methods from the present invention can be used for detecting the presence of Chlamydia in a patient; to stimulate and/or expand T cells specific for a Chlamydia protein; and for treatment of a Chlamydia infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent sequences used in the exemplification of the present invention

XX Sequence 1006 AA;

SQ Query Match 96.6%; Score 5090; DB 5; Length 1006;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 979; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 31 MVPOGIYDGETLTVSPPTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFGNLGSFT 90

Db 25 MIPOGIYDGETLTVSPPTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFGNLGSFT 84

Qy 91 VLGRHSLTPENIRTSNGAALSAAADGLFTIEGFKELSFNSCNLSLAVLPAATTNKG 150

Db 85 VLGRHSLTPENIRTSNGAALSAAADGLFTIEGFKELSFNSCNLSLAVLPAATTNKG 144

Qy 151 QTPTTSTPSNGTIYSKTDLLLNNEKFSFYNLSVGGDAIDAKSLTVQGISKLCVFOE 210

Db 145 QTPTTSTPSNGTIYSKTDLLLNNEKFSFYNLSVGGDAIDAKSLTVQGISKLCVFOE 204

Qy 211 NTAQADGACQVTSFSAANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 270

Db 205 NTAQADGACQVTSFSAANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 264

Qy 271 SFSRNTAVFPGNVARVGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 330

Db 265 SFSRNTAVFPGNVARVGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 324

Qy 331 SNNTYDGGGAI FCKNGAQAGSNNSGVSPDGRGVVFPFSSNVAAGKGAIYAKKLSVANC 390

Db 325 SNNTYDGGGAI FCKNGAQAGSNNSGVSPDGRGVVFPFSSNVAAGKGAIYAKKLSVANC 384

QY 391 VOFLRNANDGGAIVLGSSELSDADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 450  
 DB 395 VOFLRNANDGGAIVLGSSELSDADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 444  
 QY 451 GSGGKITTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510  
 DB 445 GSGGKITTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 504  
 QY 511 YQNVITIEQRIIVLREKAKLSVNSLSQTGSSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 570  
 DB 505 YQNVITIEQRIIVLREKAKLSVNSLSQTGSSLYMEAGSTLDFVTPPOPPQPPAANQLITLS 564  
 QY 571 NLHLSLSLLANNVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDTDYDRYD 630  
 DB 565 NLHLSLSLLANNVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDTDYDRYD 624  
 QY 631 WLGSNQKINVLQLQGTGPPANAPSDLTILGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 690  
 DB 625 WLGSNQKINVLQLQGTGPPANAPSDLTILGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 684  
 QY 691 TWTGTGYNPGERVASLVPNSLWGSITLDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFVH 750  
 DB 685 TWTGTGYNPGERVASLVPNSLWGSITLDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFVH 744  
 QY 751 DRDALGQGYRYISGGYSLGANSYFGSSMFLGAFTEVFGRSKDYVVCRSNHHACIGSVILS 810  
 DB 745 DRDALGQGYRYISGGYSLGANSYFGSSMFLGAFTEVFGRSKDYVVCRSNHHACIGSVILS 804  
 QY 811 TQOALCGSLYFGDAFIRASYGFGNQHMKTSTYTPAESDVRWNNCLAGHIGAGLPVITP 870  
 DB 805 TQOALCGSLYFGDAFIRASYGFGNQHMKTSTYTPAESDVRWNNCLAGHIGAGLPVITP 864  
 QY 871 SKLYLNELPFVQAEFSYADHESFTEGQOARAFKSGHLLNLSVPYGVKFDRCSSSTHPNK 930  
 DB 865 SKLYLNELPFVQAEFSYADHESFTEGQOARAFKSGHLLNLSVPYGVKFDRCSSSTHPNK 924  
 QY 931 YSPMAAYICDVAHTISGTETLLSHOETWTTDAFLARHGTVVVRGSMYASLTNSIEVYGH 990  
 DB 925 YSPMAAYICDVAHTISGTETLLSHOETWTTDAFLARHGTVVVRGSMYASLTNSIEVYGH 984  
 QY 991 GRYEYRDASRGYLSAGSRVRF 1012  
 DB 985 GRYEYRDASRGYLSAGSRVRF 1006  
 RESULT 26  
 AAB13633  
 ID AAB13633 standard; protein; 982 AA.  
 XX AAB13633;  
 AC AAB13633;  
 DT 02-FEB-2001 (first entry)  
 XX C. trachomatis pmpG gene protein.  
 DE Chlamydia infection; sexually transmitted disease;  
 XX pelvic inflammatory disease; PID; tubal obstruction; infertility;  
 KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;  
 KW coronary heart disease; antibacterial.  
 XX Chlamydia trachomatis.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 981  
 FT /note= "Unspecified amino acid"  
 FT  
 XX WO2000034483-A2.  
 XX  
 XX 15-JUN-2000.  
 PD  
 XX 08-DEC-1999; 99WO-US029012.  
 PF  
 XX 08-DEC-1998; 98US-00208277.  
 PR

PR 08-APR-1999; 99US-00288594.  
 PR 01-OCT-1999; 99US-00410568.  
 PR 22-OCT-1999; 99US-00426571.  
 XX (CORI-) CORIXA CORP.  
 PA Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
 PI WPI; 2000-431303/37.  
 XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
 PT comprises immunogenic portion of Chlamydia antigen, which comprises amino  
 PT acid sequence encoded by polynucleotide sequence.  
 XX Claim 2; Page 181-184; 256pp; English.  
 XX The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydiae are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention  
 XX Sequence 982 AA;  
 SQ Query Match 96.5%; Score 5084; DB 3; Length 982;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 31 MYPQGIYDGETLTVSPFTVIGDPSGTTVFSAGELTLKNDNSIAALPLSCFNLGSGPT 90  
 DB 1 MYPQGIYDGETLTVSPFTVIGDPSGTTVFSAGELTLKNDNSIAALPLSCFNLGSGPT 60  
 QY 91 VLGRGHSLTFFENIRTSNGAALSNAADGLFTIEGFKELSFNCNSLLAVLPAATNKG 150  
 DB 61 VLGRGHSLTFFENIRTSNGAALSNAADGLFTIEGFKELSFNCNSLLAVLPAATNKG 120  
 QY 151 QTPTTTSPNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 210  
 DB 121 QTPTTTSPNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 180  
 QY 211 NTAQADGGACQVVTSPSAMANEAPAFVANVAGVRGGIAAVQDGGQGVSSSTSTEDPVV 270  
 DB 181 NTAQADGGACQVVTSPSAMANEAPAFVANVAGVRGGIAAVQDGGQGVSSSTSTEDPVV 240  
 QY 271 SPSRNTAVFDPGNVAVRGVGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGASNT 330  
 DB 241 SPSRNTAVFDPGNVAVRGVGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGASNT 300  
 QY 331 SNNYGDDGAI FCKNGAAGSNNSGVSFDGEGVVFSSNVAAGKGAIVAKKLSVANC 390  
 DB 301 SNNYGDDGAI FCKNGAAGSNNSGVSFDGEGVVFSSNVAAGKGAIVAKKLSVANC 360  
 QY 391 VOFLRNANDGGAIVLGSSELSDADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 450  
 DB 361 VOFLRNANDGGAIVLGSSELSDADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 420  
 QY 451 GSGGKITTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510  
 DB 421 GSGGKITTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 480  
 QY 511 YQNVITIEQRIIVLREKAKLSVNSLSQTGSSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 570  
 DB 481 YQNVITIEQRIIVLREKAKLSVNSLSQTGSSLYMEAGSTLDFVTPPOPPQPPAANQLITLS 540

```
QY 571 NLHLSLLANNVNTPTPPAQDSHPAVIGTSTAGSVTISGPIFFEDLDDTAYDRYD 630
Db 541 NLHLSLLANNVNTPTPPAQDSHPAVIGTSTAGSVTISGPIFFEDLDDTAYDRYD 600
QY 631 WLGSNOKINVLKQLGKTPPANAPSDLTGEMPKYQGSWKLAWDPNTANNPPTLKA 690
Db 601 WLGSNOKINVLKQLGKTPPANAPSDLTGEMPKYQGSWKLAWDPNTANNPPTLKA 660
QY 691 TWTGTGYNPGPERVASLVPNSLWMSILDIRSAHSAIQASVDGRSYCRGLWVGVSNNFFYH 750
Db 661 TWTGTGYNPGPERVASLVPNSLWMSILDIRSAHSAIQASVDGRSYCRGLWVGVSNNFFYH 720
QY 751 DRDALGQGYRYISGYSILGANSYFGSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS 810
Db 721 DRDALGQGYRYISGYSILGANSYFGSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS 780
QY 811 TQOALCGSYLFGDAFIRASYGFGNQHMTSYTFAEESDVRDNNCLAGEIGAGLPVITP 870
Db 781 TQOALCGSYLFGDAFIRASYGFGNQHMTSYTFAEESDVRDNNCLAGEIGAGLPVITP 840
QY 871 SKLYLNEILRPVQAEFSAHESFTEEGDQARAFKSGHLLNLSPVGVKFPDRCSSTHPNK 930
Db 841 SKLYLNEILRPVQAEFSAHESFTEEGDQARAFKSGHLLNLSPVGVKFPDRCSSTHPNK 900
QY 931 YSPMAAVICDAVRTISGTETLLSHOETWTTDAFHLARHGUVVRGSMYASLTSNIEVYGH 990
Db 901 YSPMAAVICDAVRTISGTETLLSHOETWTTDAFHLARHGUVVRGSMYASLTSNIEVYGH 960
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012
Db 961 GRYEYRDASRGYGLSAGSKVAF 982

RESULT 27
AAG83201
ID AAG83201 standard; protein; 982 AA.
XX
AC AAG83201;
XX
DT 05-SEP-2001 (first entry)
XX
DE Protein encoded by Chlamydia trachomatis pmpG gene.
XX
KW Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; Cap1; CT529; OMCB;
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
OS Chlamydia trachomatis.
XX
PN WO200140474-A2.
XX
PD 07-JUN-2001.
XX
PF 04-DEC-2000; 2000WO-US032919.
XX
PR 03-DEC-1999; 99US-00454684.
PR 19-APR-2000; 2000US-00556877.
PR 20-JUN-2000; 2000US-00598419.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAM, Fling SP, Scholler J;
XX
DR WPI; 2001-374831/39.
XX
PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT inflammatory disease, trachoma, acute respiratory tract infections,
PT atherosclerosis and heart disease.
XX
PS Claim 2; Page 189-191; 295pp; English.
XX
XX The present sequence is provided in a specification relating to compounds
```

```
CC and methods for the treatment and diagnosis of chlamydial infection. The
CC compounds provided include polypeptides and fusion proteins comprising
CC immunogenic portions of Chlamydia antigens and DNA sequences encoding
CC such polypeptides. They are useful for vaccinating against chlamydial
CC infection, which causes pelvic inflammatory disease, trachoma, acute
CC respiratory tract infections, atherosclerosis and heart disease
XX
SQ Sequence 982 AA;
Query Match 96.5%; Score 5084; DB 4; Length 982;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 31 MVPGGIDGETLTVSPPYTVIGDPSTGTTVSAGBELTLKNLDNSTAAALPLSCFGLNLSFT 90
Db 1 MIPQGIYDGETLTVSPPYTVIGDPSTGTTVSAGBELTLKNLDNSTAAALPLSCFGLNLSFT 60
QY 91 VLGRGHSILPENIRTSNGAALSNSADGLFTIEGFKELSFNSCNLSLAVLPAATTNKG 150
Db 61 VLGRGHSILPENIRTSNGAALSNSADGLFTIEGFKELSFNSCNLSLAVLPAATTNKG 120
QY 151 QTPTTTTPSNGTTSYKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVFOE 210
Db 121 QTPTTTTPSNGTTSYKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVFOE 180
QY 211 NTAQADGACQVTVSPSAMANEAPIAFVANVAGVRGGIAAIVQDQGGVSSSTSTEDPVV 270
Db 181 NTAQADGACQVTVSPSAMANEAPIAFVANVAGVRGGIAAIVQDQGGVSSSTSTEDPVV 240
QY 271 SPSENTAVFPGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYTAAKOPTSGQASNT 330
Db 241 SPSENTAVFPGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYTAAKOPTSGQASNT 300
QY 331 SNNTYDGGAIKCKXGAQAGSNNSGVSFSDGEGVVFSSNVAAGGAIYAKKLSVANCGP 390
Db 301 SNNTYDGGAIKCKXGAQAGSNNSGVSFSDGEGVVFSSNVAAGGAIYAKKLSVANCGP 360
QY 391 VQFLRNIANDGAIYLGESGELSLSADYDGIIFDGNLKRKTAKENAAADVNGVTVSSQAISM 450
Db 361 VQFLRNIANDGAIYLGESGELSLSADYDGIIFDGNLKRKTAKENAAADVNGVTVSSQAISM 420
QY 451 GSGGKITTLRAKAGHOILFNDPIEMANGNNOQASSKLLKINDGEGTGDIVPANGSSTL 510
Db 421 GSGGKITTLRAKAGHOILFNDPIEMANGNNOQASSKLLKINDGEGTGDIVPANGSSTL 480
QY 511 YQNTYTIQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLS 570
Db 481 YQNTYTIQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLS 540
QY 571 NLHLSLLANNVNTPTPPAQDSHPAVIGTSTAGSVTISGPIFFEDLDDTAYDRYD 630
Db 541 NLHLSLLANNVNTPTPPAQDSHPAVIGTSTAGSVTISGPIFFEDLDDTAYDRYD 600
QY 631 WLGSNOKINVLKQLGKTPPANAPSDLTGEMPKYQGSWKLAWDPNTANNPPTLKA 690
Db 601 WLGSNOKINVLKQLGKTPPANAPSDLTGEMPKYQGSWKLAWDPNTANNPPTLKA 660
QY 691 TWTGTGYNPGPERVASLVPNSLWMSILDIRSAHSAIQASVDGRSYCRGLWVGVSNNFFYH 750
Db 661 TWTGTGYNPGPERVASLVPNSLWMSILDIRSAHSAIQASVDGRSYCRGLWVGVSNNFFYH 720
QY 751 DRDALGQGYRYISGYSILGANSYFGSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS 810
Db 721 DRDALGQGYRYISGYSILGANSYFGSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS 780
QY 811 TQOALCGSYLFGDAFIRASYGFGNQHMTSYTFAEESDVRDNNCLAGEIGAGLPVITP 870
Db 781 TQOALCGSYLFGDAFIRASYGFGNQHMTSYTFAEESDVRDNNCLAGEIGAGLPVITP 840
QY 871 SKLYLNEILRPVQAEFSAHESFTEEGDQARAFKSGHLLNLSPVGVKFPDRCSSTHPNK 930
Db 841 SKLYLNEILRPVQAEFSAHESFTEEGDQARAFKSGHLLNLSPVGVKFPDRCSSTHPNK 900
```

QY	931	YSFMAAICDAVRTISGTETLLSHQETTTTDAFHLARHGVVVRGSMVASLTSNIEVYGH	990
DB	901	YSFMAAICDAVRTISGTETLLSHQETTTTDAFHLARHGVVVRGSMVASLTSNIEVYGH	960
QY	991	GRYEYRDASRGYGLSAGSRVRF	1012
DB	961	GRYEYRDASRGYGLSAGSKVXF	982
RESULT 28			
ID	ABB94172	standard; protein; 982 AA.	
XX	AC	ABB94172;	
XX	05-JUN-2002	(first entry)	
DE	DE	Chlamydia protein sequence SEQ ID NO:176.	
XX	KW	Chlamydia infection; Chlamydia; vaccine; detection; diagnosis; antigen;	
KW	antibacterial; immunostimulant; immune response;		
KW	Chlamydia-specific T-cell response.		
XX	OS	Chlamydia sp.	
XX	WO2002028267-A2.		
PN	31-JAN-2002.		
PD	20-JUL-2001;	2001WO-US023121.	
XX	20-JUL-2000;	2000US-00620412.	
PR	23-APR-2001;	2001US-00841132.	
XX	(CORI-) CORIXA CORP.		
XX	PI	Fling SP, Skeiky YAW, Probst P, Bhatia A;	
XX	WPI;	2002-179901/23.	
XX	Novel compositions comprising Chlamydia Cap1 protein and its use in the		
PT	treatment of Chlamydia infection.		
XX	Disclosure; Page 221-223; 537pp; English.		
XX	The present invention describes compositions comprising a Chlamydia Cap1		
CC	protein and methods for the diagnosis and therapy of Chlamydia infection.		
CC	Chlamydia DNA and protein sequences from the present invention can have		
CC	antibacterial and immunostimulant activities, and can be used in		
CC	vaccines. Compounds from the present invention can be used for eliciting		
CC	an immune response, specifically stimulating a Chlamydia-specific T-cell		
CC	response or inhibiting the development of a Chlamydia infection in an		
CC	animal. Methods from the present invention can be used for detecting the		
CC	presence of Chlamydia in a patient; to stimulate and/or expand T cells		
CC	specific for a Chlamydia protein; and for treatment of a Chlamydia		
CC	infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent		
CC	sequences used in the exemplification of the present invention		
XX	Sequence 982 AA;		
SQ	Query Match	96.5%; Score 5084; DB 5; Length 982;	
	Best Local Similarity	99.6%; Pred. No. 0;	
	Matches 978; Conservative	2; Mismatches 2; Indels 0; Gaps 0;	
QY	31	MVPGIYDGETLTVSRPVTYVIGDPSTVFSAGELTKNLDNSIAALPLSCFNGLLGSFT	90
DB	1	MIPQIYDGETLTVSRPVTYVIGDPSTVFSAGELTKNLDNSIAALPLSCFNGLLGSFT	60
QY	91	VLGRGHSLTFTENRTSTNGAALSNSAADGLTIEGPKELSFNSCNSLLAVLPAATTNKG	150
DB	61	VLGRGHSLTFTENRTSTNGAALSNSAADGLTIEGPKELSFNSCNSLLAVLPAATTNKG	120
QY	151	QTPTTSTPSNGTIYSKTDLLLLNNEKFSFYSLNLSVSGDGAIDAKSLTVQGISKLCVQFE	210

DB	121	QTPTTSTPSNGTIYSKTDLLLLNNEKFSFYSLNLSVSGDGAIDAKSLTVQGISKLCVQFE	180
QY	211	NTAQDGGACQCVVTSFSAMANEAPAFVANVAGVRGGGIAAVQDQCGQCVSSSTSTEDPV	270
DB	181	NTAQDGGACQCVVTSFSAMANEAPAFVANVAGVRGGGIAAVQDQCGQCVSSSTSTEDPV	240
QY	271	SFSRNTAVEFDGNVARVGGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNT	330
DB	241	SFSRNTAVEFDGNVARVGGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNT	300
QY	331	SNNYDGGAIIFCKGQAQAGSNNSGVSFDPGCVVPFSSNVAAGKGAIAKLSVANGCP	390
DB	301	SNNYDGGAIIFCKGQAQAGSNNSGVSFDPGCVVPFSSNVAAGKGAIAKLSVANGCP	360
QY	391	VOFLRNANDGGAIVLGESGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAISM	450
DB	361	VOFLRNANDGGAIVLGESGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAISM	420
QY	451	CSGGKITTLRAKAGHQIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIYFANGSSTL	510
DB	421	CSGGKITTLRAKAGHQIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIYFANGSSTL	480
QY	511	YONVTIEQGRIVLREKAKLSVNSLSQTCGSLYMEAGSTWDFVTPPOPPPAANQLITLS	570
DB	481	YONVTIEQGRIVLREKAKLSVNSLSQTCGSLYMEAGSTWDFVTPPOPPPAANQLITLS	540
QY	571	NLHLSLLANNVTPNPPAQQDHPAVTIGSTTAGSVTISGPIFPEDLDLDTAYDRYD	630
DB	541	NLHLSLLANNVTPNPPAQQDHPAVTIGSTTAGSVTISGPIFPEDLDLDTAYDRYD	600
QY	631	WLGSNQKINVLKQLGTGKPPANAPSDLTIGNEMPKYGYQGSWKLAWDPNTANNPPTLKA	690
DB	601	WLGSNQKINVLKQLGTGKPPANAPSDLTIGNEMPKYGYQGSWKLAWDPNTANNPPTLKA	660
QY	691	TWTKGTGYPGPVRSVSLVNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	750
DB	661	TWTKGTGYPGPVRSVSLVNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	720
QY	751	DRDALGQGYRYISGGYSLGANSYFGSSMFGLAFTVEVGRSKDYVVCRRSHHACIGSVYLS	810
DB	721	DRDALGQGYRYISGGYSLGANSYFGSSMFGLAFTVEVGRSKDYVVCRRSHHACIGSVYLS	780
QY	811	TQOALCGSYLFGDAFIRASYFGNQHMKTSYTFABESDVRNDNNCLAGEIGAGLPVITP	870
DB	781	TQOALCGSYLFGDAFIRASYFGNQHMKTSYTFABESDVRNDNNCLAGEIGAGLPVITP	840
QY	871	SKLYLNLRLPFFVQAEFSYADHESFTEEGDQARAFKSGHLLNLSPVGVKFDRCSTHPNK	930
DB	841	SKLYLNLRLPFFVQAEFSYADHESFTEEGDQARAFKSGHLLNLSPVGVKFDRCSTHPNK	900
QY	931	YSFMAAICDAVRTISGTETLLSHQETTTTDAFHLARHGVVVRGSMVASLTSNIEVYGH	990
DB	901	YSFMAAICDAVRTISGTETLLSHQETTTTDAFHLARHGVVVRGSMVASLTSNIEVYGH	960
QY	991	GRYEYRDASRGYGLSAGSRVRF	1012
DB	961	GRYEYRDASRGYGLSAGSKVXF	982

RESULT 29			
ID	AAB13639	standard; protein; 1006 AA.	
XX	AC	AAB13639;	
XX	02-FEB-2001	(first entry)	
XX	DE	C. trachomatis pmpG gene protein.	
XX	KW	Chlamydial infection; sexually transmitted disease;	
KW	pelvic inflammatory disease; PID; tubal obstruction; infertility;		
KW	trachoma; blindness; acute respiratory tract infection; atherosclerosis;		



KW coronary heart disease; antibacterial.  
XX Chlamydia trachomatis.  
OS WO200034483-A2.  
XX 15-JUN-2000.  
XX 08-DEC-1999; 99WO-US029012.  
XX 08-DEC-1998; 98US-00208277.  
PR 08-APR-1999; 99US-00288594.  
PR 01-OCT-1999; 99US-00410568.  
PR 22-OCT-1999; 99US-00426571.  
XX (CORI-) CORIXA CORP.  
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;  
XX WPI, 2000-431303/37.  
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
PT comprises immunogenic portion of Chlamydia antigen, which comprises amino  
PT acid sequence encoded by polynucleotide sequence.  
XX Claim 2; Page 208-210; 256pp; English.  
XX The present invention relates to new nucleic acid sequences and the  
CC proteins encoded by the nucleic acid sequences. The encoded proteins  
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
CC infection. Chlamydiae are intracellular bacterial pathogens that are  
CC responsible for a wide variety of human infections. C. trachomatis  
CC infection is one of the most common sexually transmitted diseases and  
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
CC the leading cause of preventable blindness worldwide. C. pneumoniae is a  
CC major cause of acute respiratory tract infections in humans and is also  
CC thought to play a role in the pathogenesis of atherosclerosis and  
CC coronary heart disease. The present invention is a protein isolated in the  
XX present invention  
XX Sequence 1006 AA;  
SQ

Query Match 96.5%; Score 5083; DB 3; Length 1006;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 31 MVRQGIYDGETLTWSPPYTVIGDPBGTTFPSAGELTLKNLNSIAALPLSCFNLGSPT 90  
DB 25 MIPQGIYDGETLTWSPPYTVIGDPBGTTFPSAGELTLKNLNSIAALPLSCFNLGSPT 84  
QY 91 VLGRGHSITPENIRTSNGAALSNSAADGLFTIEGPKELSPNSCNLSLLAVLPAATTNKG 150  
DB 85 VLGRGHSITPENIRTSNGAALSNSAADGLFTIEGPKELSPNSCNLSLLAVLPAATTNKG 144  
QY 151 QTPTTSTPSNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTVQGISKLCVFOE 210  
DB 145 QTPTTSTPSNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTVQGISKLCVFOE 204  
QY 211 NTAQADGACQVTFSPSMAANEAFIAFVANVAGVGGIAAVQDQOQVSSSTEDPVV 270  
DB 205 NTAQADGACQVTFSPSMAANEAFIAFVANVAGVGGIAAVQDQOQVSSSTEDPVV 264  
QY 271 SFSRNTAVEPDGNVARGGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQASNT 330  
DB 265 SFSRNTAVEPDGNVARGGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQASNT 324  
QY 331 SNNTYGDGGAIFCKNGAQAGSNNSGVSFDPGEGVVPFSSNVAAGKGGAIYAKKLSVANGCP 390  
DB 325 SNNTYGDGGAIFCKNGAQAGSNNSGVSFDPGEGVVPFSSNVAAGKGGAIYAKKLSVANGCP 384  
QY 391 VOFLRNANDGGAIYLGESGELSADYGDIIIPGNLKRKTAKENADVNGVTVSSQAISM 450

DB 385 VOFLRNANDGGAIYLGESGELSADYGDIIIPGNLKRKTAKENADVNGVTVSSQAISM 444  
QY 451 GSGGKITTLRAKAGHQLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 510  
DB 445 GSGGKITTLRAKAGHQLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 504  
QY 511 YQNTVIEGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPQPPQPPAAQLITLS 570  
DB 505 YQNTVIEGRIVLREKAKLSVNSLSQSGSLYMEAGSTLDFVTPQPPQPPAAQLITLS 564  
QY 571 NLHLSLSLLANNAVTPPTNPADSDHPAIVIGSTTAGSVTISGPIFFEDLDDTAYDRYD 630  
DB 565 NLHLSLSLLANNAVTPPTNPADSDHPAIVIGSTTAGSVTISGPIFFEDLDDTAYDRYD 624  
QY 631 WLGSNOKINVLKQLQGTKPPANAPSDLTGLNEMPKYQSGSKLAWDPNTANNNGPYTLKA 690  
DB 625 WLGSNOKINVLKQLQGTKPPANAPSDLTGLNEMPKYQSGSKLAWDPNTANNNGPYTLKA 684  
QY 691 TWTGTGNPGPERSVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFPYH 750  
DB 685 TWTGTGNPGPERSVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFPYH 744  
QY 751 DRDALGOGYRYISGGYSLGANSYFGSSMFGLAFTFVGRSKDYVYCRSNHHACIGSVYLS 810  
DB 745 DRDALGOGYRYISGGYSLGANSYFGSSMFGLAFTFVGRSKDYVYCRSNHHACIGSVYLS 804  
QY 811 TQALCCSYLFGDAFIRASYGFGNQMKTSTYFAESDVRMDNCLAGEIGAGLPIVITP 870  
DB 805 TQALCCSYLFGDAFIRASYGFGNQMKTSTYFAESDVRMDNCLAGEIGAGLPIVITP 864  
QY 871 SKLYNLRLPFVQAEFSAADHESFTERGDOARAPKSGHLLNLSPVGVGVKFDRCSTHPNK 930  
DB 865 SKLYNLRLPFVQAEFSAADHESFTERGDOARAPKSGHLLNLSPVGVGVKFDRCSTHPNK 924  
QY 931 YSFMAAVICDAYRTISGTETTLTSHQETWTTDAPLHARHGTVVGRGSMYASLTSNIEVYGH 990  
DB 925 YSFMAAVICDAYRTISGTETTLTSHQETWTTDAPLHARHGTVVGRGSMYASLTSNIEVYGH 984  
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012  
DB 985 GRYEYRDASRGYGLSAGSKVRF 1006  
RESULT 30  
ADD42756  
ID ADD42756 standard; protein; 670 AA.  
XX AC ADD42756;  
XX DT 15-JAN-2004 (first entry)  
XX Chlamydia pmpG passenger domain protein SEQ ID NO:169.  
XX Chlamydia infection; Chlamydia; antibiotic; antiinflammatory;  
XX antifertility; cardant; antiarteriosclerotic; ophthalmological;  
XX vaccine; gene therapy; immune response; pelvic inflammatory disease;  
XX tubal obstruction; infertility; male infertility; ocular infection;  
XX blindness; acute respiratory tract infection; atherosclerosis;  
XX coronary heart disease.  
XX Chlamydia sp.  
XX OS WO2003041560-A2.  
XX FN 22-MAY-2003.  
XX PD 05-NOV-2002; 2002WO-US035624.  
XX PF 06-NOV-2001; 2001US-00012256.  
XX PR 05-DEC-2001; 2001US-00007693.  
XX PR 15-JUL-2002; 2002US-00197220.  
XX



QY 142 PAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQG 201  
 Db 63 PAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQG 122  
 QY 202 ISKLCVFQNTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSS 261  
 Db 123 ISKLCVFQNTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSS 182  
 QY 262 STSTEDPVVFSRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAQ 321  
 Db 183 STSTEDPVVFSRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAQ 242  
 QY 322 PTSQASNTSNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAK 381  
 Db 243 PTSQASNTSNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAK 302  
 QY 382 KLSVANCGPVFLRNIANDGGAIVLGESELSADYDIIIFDGNLKRKTAKENADVNGV 441  
 Db 303 KLSVANCGPVFLRNIANDGGAIVLGESELSADYDIIIFDGNLKRKTAKENADVNGV 362  
 QY 442 TVSSQASISMGSGGKITTLRAKAGHQILFNDPIEMANGNQPQASKLLKINDGEGYTDI 501  
 Db 363 TVSSQASISMGSGGKITTLRAKAGHQILFNDPIEMANGNQPQASKLLKINDGEGYTDI 422  
 QY 502 VPANGSSTLYQNVITBOGRIVLREKAKLSVNSLSQTGSLYMBAGSTWDFVTPPPQOPP 561  
 Db 423 VPANGSSTLYQNVITBOGRIVLREKAKLSVNSLSQTGSLYMBAGSTWDFVTPPPQOPP 482  
 QY 562 AANQLITLSNLSLSLLANNAVTPNPTNPPAQDS 597  
 Db 483 AANQLITLSNLSLSLLANNAVTPNPTNPPAQDS 518

RESULT 32

AA16739  
 ID AA16739 standard; protein; 505 AA.  
 XX  
 AC AA16739;  
 XX  
 DT 21-JUL-1999 (first entry)  
 XX  
 DE C. trachomatis L2 HMW protein fragment.  
 XX  
 KW Chlamydia; high molecular weight protein; HMW protein; urethritis;  
 KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;  
 KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;  
 KW salpingitis; tubal occlusion; infertility; cervical cancer;  
 KW arteriosclerosis; atherosclerosis.  
 XX  
 OS Chlamydia trachomatis.  
 XX  
 FN WO9917741-A1.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 01-OCT-1998; 98MO-US020737.  
 XX  
 PR 02-OCT-1997; 97US-00942596.  
 XX  
 PA (ANTE-) ANTEX BIOLOGICS INC.  
 XX  
 PI Jackson JW, Pace JL;  
 XX  
 DR WPI; 1999-287659/24.  
 XX  
 XX New Chlamydia protein useful for treating conjunctivitis, urethritis and  
 PT cervical cancer.  
 XX  
 XX Claim 5; Page 123-125; 141pp; English.  
 XX  
 XX The invention relates to an isolated Chlamydia species high molecular  
 CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as

CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can  
 CC be used for preventing, treating or ameliorating a disorder related to  
 CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,  
 CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,  
 CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical  
 CC cancer, infertility, arteriosclerosis and atherosclerosis. The products  
 CC can also be used for detection and diagnosis. The present sequence  
 CC represents a C. trachomatis HMW protein fragment  
 XX

SQ Sequence 505 AA;

Query Match 48.4%; Score 2547; DB 2; Length 505;  
 Best Local Similarity 99.8%; Pred. No. 2.4e-167;  
 Matches 504; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 29 EIMVPOGIYDGETLTVSPFTVIGDPGSGTTFVFSAGELTLKLNLSIAALPLSCFNLGGS 88  
 Db 1 EIMVPOGIYDGETLTVSPFTVIGDPGSGTTFVFSAGELTLKLNLSIAALPLSCFNLGGS 60  
 QY 89 FTVLGRGHSITFENIRTSNKAALNSAADGLFTIEGKELSPNSCNLSLAVLPAATTNK 148  
 Db 61 FTVLGRGHSITFENIRTSNKAALNSAADGLFTIEGKELSPNSCNLSLAVLPAATTNK 120  
 QY 149 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQGISKLCVP 208  
 Db 121 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQGISKLCVP 180  
 QY 209 QENTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDP 268  
 Db 181 QENTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDP 240  
 QY 269 VVSFRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAQPTSGQAS 328  
 Db 241 VVSFRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAQPTSGQAS 300  
 QY 329 NTSNNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAKLSVANC 388  
 Db 301 NTSNNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAKLSVANC 360  
 QY 389 GPVQFLRNIANDGGAIVLGESELSADYDIIIFDGNLKRKTAKENADVNGVTVSSQAI 448  
 Db 361 GPVQFLRNIANDGGAIVLGESELSADYDIIIFDGNLKRKTAKENADVNGVTVSSQAI 420  
 QY 449 SMGSGGKITTLRAKAGHQILFNDPIEMANGNQPQASKLLKINDGEGYTDIVFANGSS 508  
 Db 421 SMGSGGKITTLRAKAGHQILFNDPIEMANGNQPQASKLLKINDGEGYTDIVFANGSS 480  
 QY 509 TLQNVTIEQGRIVLREKAKLSVNS 533  
 Db 481 TLQNVTIEQGRIVLREKAKLSVNS 505

RESULT 33

AA16751  
 ID AA16751 standard; protein; 458 AA.  
 XX  
 AC AA16751;  
 XX  
 DT 21-JUL-1999 (first entry)  
 XX  
 DE Chlamydia HMW protein fragment.  
 XX  
 KW Chlamydia; high molecular weight protein; HMW protein; urethritis;  
 KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;  
 KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;  
 KW salpingitis; tubal occlusion; infertility; cervical cancer;  
 KW arteriosclerosis; atherosclerosis.  
 XX  
 OS Chlamydia sp.  
 XX  
 FN WO9917741-A1.  
 XX  
 PD 15-APR-1999.

```

XX 01-OCT-1998; 98WO-US020737.
XX
XX 02-OCT-1997; 97US-00942596.
XX
XX (ANTE-) ANTEX BIOLOGICS INC.
XX
XX Jackson JW, Pace JL;
XX WPI; 1999-287659/24.
XX
XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
XX cervical cancer.
XX
XX Claim 5; Page 133-135; 141pp; English.
XX
XX The invention relates to an isolated Chlamydia species high molecular
XX weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
XX determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
XX be used for preventing, treating or ameliorating a disorder related to
XX Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
XX lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
XX pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
XX cancer, infertility, arteriosclerosis and atherosclerosis. The products
XX can also be used for detection and diagnosis. Sequences AAY16740-Y16752
XX represent Chlamydia HMW protein fragments
XX
XX Sequence 458 AA;
XX
XX Query Match 44.6%; Score 2350; DB 2; Length 458;
XX Best Local Similarity 100.0%; Pred. No. 9e-154;
XX Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 217 GGACQVVTFSAMANEAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVVFSRNT 276
DB 1 GGACQVVTFSAMANEAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVVFSRNT 60
QY 277 AVEFDGNVAVGGIYSGNVAFNNGKTLFLNNVAVPVYIAAKQPTSGOASNTSNYGD 336
DB 61 AVEFDGNVAVGGIYSGNVAFNNGKTLFLNNVAVPVYIAAKQPTSGOASNTSNYGD 120
QY 337 GGAIFCKNGAAGSNNSGVSFDGEGVWPFSSNVAAGKGAIYAKKLSVANCGPVQFLRN 396
DB 121 GGAIFCKNGAAGSNNSGVSFDGEGVWPFSSNVAAGKGAIYAKKLSVANCGPVQFLRN 180
QY 397 IANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENADVNGVTVSSQAIISMGSGGKI 456
DB 181 IANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENADVNGVTVSSQAIISMGSGGKI 240
QY 457 TTLRAKAGHQLFNDPIEMANGNQPAAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 516
DB 241 TTLRAKAGHQLFNDPIEMANGNQPAAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 300
QY 517 EQGRIVLREKAKLSVNSLSOTGSLYMEAGSTWDFVTPQPQPPPAANOLITLSNLHLSL 576
DB 301 EQGRIVLREKAKLSVNSLSOTGSLYMEAGSTWDFVTPQPQPPPAANOLITLSNLHLSL 360
QY 577 SLLANNAVTPPTNPPAOSHFAVIGSTTAGSVTISGPIFFBDLDDTAYDRYDWLGSNQ 636
DB 361 SLLANNAVTPPTNPPAOSHFAVIGSTTAGSVTISGPIFFBDLDDTAYDRYDWLGSNQ 420
QY 637 KINVLKQLGTQKPPANAPSDLTIGNEMPKYGYQGSWKL 674
DB 421 KINVLKQLGTQKPPANAPSDLTIGNEMPKYGYQGSWKL 458
XX
XX RESULT 34
XX ID AAY16752
XX AC AAY16752;
XX XX 21-JUL-1999 (first entry)

```

```

XX Chlamydia HMW protein fragment.
XX
XX Chlamydia; high molecular weight protein; HMW protein; urethritis;
XX bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
XX cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
XX salpingitis; tubal occlusion; infertility; cervical cancer;
XX arteriosclerosis; atherosclerosis.
XX
XX Chlamydia sp.
XX
XX WO9917741-A1.
XX
XX 15-APR-1999.
XX
XX 01-OCT-1998; 98WO-US020737.
XX
XX 02-OCT-1997; 97US-00942596.
XX
XX (ANTE-) ANTEX BIOLOGICS INC.
XX
XX Jackson JW, Pace JL;
XX WPI; 1999-287659/24.
XX
XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
XX cervical cancer.
XX
XX Claim 5; Page 135-136; 141pp; English.
XX
XX The invention relates to an isolated Chlamydia species high molecular
XX weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
XX determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
XX be used for preventing, treating or ameliorating a disorder related to
XX Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
XX lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
XX pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
XX cancer, infertility, arteriosclerosis and atherosclerosis. The products
XX can also be used for detection and diagnosis. Sequences AAY16740-Y16752
XX represent Chlamydia HMW protein fragments
XX
XX Sequence 325 AA;
XX
XX Query Match 32.9%; Score 1735; DB 2; Length 325;
XX Best Local Similarity 100.0%; Pred. No. 2e-111;
XX Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 688 LKATWTKTYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSVNF 747
DB 1 LKATWTKTYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSVNF 60
QY 748 FTHDRDALCGGYRYISGGYSLGANSYFGSSMFLAFTEVFGSKDYVVCRRNHHACIGSV 807
DB 61 FTHDRDALCGGYRYISGGYSLGANSYFGSSMFLAFTEVFGSKDYVVCRRNHHACIGSV 120
QY 808 YLSTQALCGSLYFGDAFIRASVFGNQHMKTSTYTPAESDVRWNNCLAGEIGAGLPIV 867
DB 121 YLSTQALCGSLYFGDAFIRASVFGNQHMKTSTYTPAESDVRWNNCLAGEIGAGLPIV 180
QY 868 ITPSKLYLNLRLPFVQAERFSYADHESFTEGDAQARAFKSGHLLNLSPVGVKDFCRCSSTH 927
DB 181 ITPSKLYLNLRLPFVQAERFSYADHESFTEGDAQARAFKSGHLLNLSPVGVKDFCRCSSTH 240
QY 928 PNKYSFMAAYICDAYRTISGTETILLSHQETTTDAFLARHGQVVRGSMYASLTNSIEV 987
DB 241 PNKYSFMAAYICDAYRTISGTETILLSHQETTTDAFLARHGQVVRGSMYASLTNSIEV 300
XX
XX 988 YGHGRYEYRDASRGYGLSAGSRVRF 1012
XX 301 YGHGRYEYRDASRGYGLSAGSRVRF 325
XX
XX RESULT 35

```









XX	WPI, 1999-105610/09.	QY	577	SSLANNNAVTPNPPAQQDHPAVIGSTTAG-SVTISGPIFFBDDLTADVRDVLGSGN	635
DR	N-PSDB; AAX06821.	Db	491	DSLQGTN-----KAIKATAASKDVALSGPIMLVDAQGNYYEHHN-LSQQ	534
XX	Species-specific test for identifying mammals infected with Chlamydia	QY	636	QKINVLKQL-GTTPPANAPSDLTGLNEMPKYQGSWKLAW-DPNTANNPGYTLKATWT	693
PT	pneumoniae - comprises detecting antibodies specific for outer membrane	Db	535	QVFFLIELSAQGTMTTDDIP-DTPIILNTNHYGGTGTIIIVVDDATAKTNNAIL--TWT	591
PT	proteins of C. pneumoniae or nucleic acids encoding these proteins.	QY	694	KTGYNPGPVERVASLVPNSLWGSILDIRSAHSAIQAQSVDRGSRVCRGLWVSGVSNFFYHDDR	753
XX	Claim 7; Page 56-58; 115pp; English.	Db	592	KTGYKPNPERQGPLVPNSLWGSFVDVRSIQSLMDRSTSSLSSTNLWVSGIADFLHEDQK	651
XX	This polypeptide comprises the novel 96.7 kDa surface exposed protein	QY	754	ALGOQYRIYSGYSLGANSYFGS-SMFLGATFEVPGRSKDYVCKSNHHACIGSV---YL	809
CC	Omp9 of the human respiratory pathogen Chlamydia pneumoniae. Its amino	Db	652	QNQRSHSSAGYALGGGFTASENFNFAFCQLFGYDKDLKLVAKHVVVAGAMSRYRHL	711
CC	acid sequence was deduced from DNA (see AAX06821) isolated from a C.	QY	810	STQOALCQSYLFGDA----FI---RASYGFGNQHMKTSYTFABESDVWDNNCLAGEIGA	862
CC	pneumoniae expression library. The invention provides 12 novel surface	Db	712	GESKTL-AKILSGNSDSLPFFVFNARFAYGHTDNNNTTKYGVSPVKSGWGNDAFGIECGG	770
CC	exposed proteins, Omp4-Omp15 (see AAW88417-28), and nucleic acid	QY	863	GLPIVITPSKLYLNLRLPRFVQAEFSYADHESFTEERGQARAFKSGHLLNLSPVGVYKDFR	922
CC	sequences encoding them (see AAX06816-27). A new species specific test is	Db	771	AIPVVASGRSGWVDHTTFLNLEMIYAHQNDKENGTEGRSFQSEDLFNLAIVPGIKPEK	830
CC	claimed that is used to identify mammals (including humans) infected with	QY	923	CSSTHPNKYSFMAAYICDAYRTISGTETTLTSHOETWTTDAFHLARHGTVVVRGSMVASLT	982
CC	Chlamydia pneumoniae. The test comprises detecting antibodies specific	Db	831	FSD--KSTYDLSIAVVPDIRNDPGCTTTLMVSGDSWSTCGTSLSRQALLVRAGNHAPA	888
CC	for Omp4-Omp15 or detecting nucleic acid fragments encoding these outer	QY	983	SNIEVYHGRVYRDASRGYGLSAGSRVRF	1012
CC	membrane proteins, especially by PCR. The proteins are also used in the	Db	889	SNFVFQSFVEVLRGSRSSYALDLGRFRGP	918
CC	diagnosis of C. pneumoniae infection in mammals. The nucleic acids and	QY	AA94327	standard; protein; 928 AA.	
CC	proteins can also be used in the immunization of mammals, the nucleic	Db	AA94327		
CC	acids being particularly useful as DNA vaccines for effecting in vivo	QY	12-SEP-2003	(revised)	
CC	expression of antigens. The vaccines may also prevent atherosclerosis and	Db	11-AUG-2000	(first entry)	
CC	bronchial asthma, which are possibly associated with C. pneumoniae.	QY	Chlamydia pneumoniae	98kD putative outer membrane protein.	
CC	(Updated on 17-OCT-2003 to standardise OS field)	Db	Chlamydia	antigen; vaccine; infection; outer membrane protein.	
XX	Sequence 918 AA;	QY	Chlamydia pneumoniae		
XX	Query Match	QY	Chlamydia pneumoniae		
XX	Best Local Similarity 21.7%; Score 1143; DB 2; Length 918;	QY	Chlamydia pneumoniae		
XX	Mismatches 324; Conservative 163; Mismatches 393; Indels 170; Gaps 34;	QY	Chlamydia pneumoniae		
QY	1 MOTSFHKFPLSMILAYSCSLNGGGYAAEIMVPPQGIYDGETLTVSP-PYTVIGDSGTTV	QY	59		
Db	1 MRSSFLLIISLAPLLMSVSAADAALTGLSRDSYNGDTSTTEPTKAAATSDASGTTY	Db	60		
QY	60 F8AGELTTLKNLONSIAALPLCPGNLLGSFTVLGRGHSITFENIRST-NGAALNSRAD	QY	118		
Db	61 ILDGDSVISQAGKQ-TSLTTSCTFNTAGNLTLFNGFSLFHDNFISSIVAGVVSNTAAS	Db	119		
QY	119 GLPTIEGPKELSPNCNLSLAVLP-AATTNKGSTPTTTSPNGSIYKTDLLILNNEK	QY	177		
Db	120 GITKPSGF-----STLRMLAAPRTTKGAKIT-----DGLVPSIGNLDQN---	Db	161		
QY	178 F8FYNLSDGCGAIDAKSILTVQGISKLCPVQENTAQADGACQVVTFSFMANEAPIAF	QY	237		
Db	162 -----ENASSENGAINTKILSLTGSTRF-----VAF	Db	188		
QY	238 VANVAGVRGGIAAODQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGIYSGNV	QY	297		
Db	189 LGNSSSQGGAIYASGD-----SVISENAGILSFGNNSA-----TTSGGAIASGNL	Db	235		
QY	298 AFLNKGTLPLNNVSPVYIAAQPTSGQASNTSNVYGDGGAIFCKNGAQAGNNSGSVS	QY	357		
Db	236 VISNNQNIFF-----DGCKATT-----NGGAIIDCN---KAGANPPDILT	Db	271		
QY	358 FDCEGVVFFSSNVAAGKGAIYAKLSVAN-CGPVQFRLNIAND-----GGAIYLGESGEL	QY	412		
Db	272 LSGNESLHFLNNTAGNSGGAIYTKLVLSRGGVLFNNKAAANATPKGGAIALDSGEI	Db	331		
QY	413 SLSADYGDIIIFDGNLKRKTAKENADVNGVTVSSQAIMSGGGKITTIRAKAGHQIILNDP	QY	472		
Db	332 SISADLGNIIIFEGNTTSTTSGSPA-----SVTRNAIDLASNAKFLNLRATGNKVIYDP	Db	385		
QY	473 IEWANGNQAQSKLLKINDGEG--YTGDIVFA-----NGSSTLYQNTTIE	QY	517		
Db	386 ITSSGATDK-----LSLNKADAGSGNTYEGYIVFSGEKLSEELKCPDLNLSKSTTQAVELA	Db	441		
QY	518 QGRIVLRKAKLSVNSLSQTGGS-LYNEAGSTWDFVTPPPQPPQPPAANQLITLSNLHSL	QY	576		
Db	442 AGALVLKDGTVTVANTITQEGSKVMDGGTTFE-----ASAEGVTILNGLAINI	Db	490		

RESULT 39  
 AAY94327  
 ID AAY94327 standard; protein; 928 AA.  
 XX AAY94327;  
 AC AAY94327;  
 XX 12-SEP-2003 (revised)  
 DT 11-AUG-2000 (first entry)  
 XX Chlamydia pneumoniae 98kD putative outer membrane protein.  
 XX Chlamydia; antigen; vaccine; infection; outer membrane protein.  
 XX Chlamydia pneumoniae.  
 OS WO200026237-A2.  
 PN 11-MAY-2000.  
 PD 29-OCT-1999; 99WO-GB003579.  
 PF 29-OCT-1998; 98US-0106070P.  
 PR 01-MAR-1999; 99US-0122066P.  
 PR 27-OCT-1999; 99US-00428122.  
 XX (CONN-) CONNAUGHT LAB LTD.  
 XX Murdin AD, Oomen RP, Dunn PL;  
 XX WPI; 2000-365569/31.  
 DR N-PSDB; AAX27021.  
 XX Novel Chlamydia 98 kDa putative outer membrane protein antigen, used for  
 PT vaccination and protection against Chlamydia infection.  
 XX Claim 6; Fig 1; 93pp; English.  
 XX The present sequence is the 98kDa putative outer membrane protein from  
 CC Chlamydia pneumoniae. The genomic sequence was amplified using two PCR  
 CC primers. The 5' primer contains a NotI restriction site, a ribosome

CC binding site, an initiation codon and a sequence close to the 5' end of  
 CC the 98kDa putative outer membrane protein coding sequence. The 3' primer  
 CC contains the sequence encoding the C-terminal sequence of the putative  
 CC outer membrane protein and a BglI restriction site. The stop codon was  
 CC excluded and an additional nucleotide was inserted to obtain an in-frame  
 CC C-terminal fusion with the Histidine tag. The PCR product was cloned into  
 CC a eukaryotic expression vector (pCA-Myc-His) by restricting both the  
 CC vector and the PCR product with NotI and BamHI and performing a ligation  
 CC reaction. This expression vector was injected intramuscularly and  
 CC intranasally into mice, which were subsequently inoculated with Chlamydia  
 CC pneumoniae. The chlamydial lung titers of the immunised mice were lower  
 CC than those of the controls. Thus the 98kDa putative outer membrane  
 CC protein can be used as a vaccine to provide protection against Chlamydia  
 CC infections, especially Chlamydia pneumoniae infections. The present  
 CC polypeptide may also be administered orally to treat Chlamydia infection.  
 CC (Updated on 12-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 928 AA;

Query Match 21.4%; Score 1128; DB 3; Length 928;  
 Best Local Similarity 29.2%; Pred. No. 1e-68;  
 Matches 311; Conservative 169; Mismatches 394; Indels 192; Gaps 32;

QY 1 MOTSPKFFLSMILAYSCSLNGGGYAAE-IMVPGIYDGETLTVSPFVTIVGDPSTTV 59  
 DB 1 MKSSFPKVFSTPAIFPLSMI-----ATETVLDSASFQGNK-NGNFSVRSEODAGITY 54  
 QY 60 FSAGELTKLNLDNSIAALPLSCFNGLLGSFTVLGRHSLTPTNIRTST-NGAALNSAAD 118  
 DB 55 LPKGNVTLENIPTGTATKSCFNNTKGDLTPTGNGSLLFTQVDAGTVAGAAVNSSVVD 114  
 QY 119 GLFTIEGKELSPSCNLSLAVLPAATNKGSTPTTTSTPNSNGTIYSKTDLLLNNEKF 178  
 DB 115 KSTTFIGFSSLSF-----IASPGSITTKGAVSCSTGS-----LSLTKNYSL 157  
 QY 179 SPYSLNLSGDGAIDAKSLTVQIGSKLCVFBQNTAQADGACQVVTSPSAMANEAPIAV 238  
 DB 158 LPSKNFTDNGAIYAKTSLTGTMTSALFSENTSKKGAIQTSDAUTITGNQGEVSF- 216  
 QY 239 ANVAGVROGGIAAVODGQGGSSSTSTEDPVVFSRNTAVFPGNVARVGGIYSYGNVA 298  
 DB 217 -----SDNTSSDS-----GAAIFTEASVT 235  
 QY 299 FLNNGKTLFLNNAVSPVIAAKQPTSGQASNTSNNGYDGGALFCKNGAQAQSNNGSVSP 358  
 DB 236 ISNNAKVGFIDN-----KVTGASSSTTGDN-SGGAICAYK-----TSTDTKVTL 278  
 QY 359 DQGVVFPSSNVAAGKGAIVAKLSVANGCPVQFLRNIA-----DGAIVLGSSELS 413  
 DB 279 TGNQMLLFNSNTSTTGAAGIYVKLELAGGLTLFSRNSVNGGTAPKGAIAIEDSGELS 338  
 QY 414 LSADYGDIIIPGNLKRATKENAADVNGVTVSQAISMGSGGKITTLRAKAGHQLFNDDPI 473  
 DB 339 LSADSGDIVPLGN---TVTSTTPGTN-----RSSIDLGTSAKMTALRSAGRAIVFYDPI 390  
 QY 474 EWANGNPQASSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNVT 515  
 DB 391 TTGSSTT---VTDVLKVNTPADSAQYTGNIIFTGKLSBTEAADSKNLTSKLLQPV 446  
 QY 516 TEQGRIVLREKAKLSVNSLSOTGGS-LYMEAGSTWDFVTPOPPQPPAANQLITLSNHL 574  
 DB 447 LSGGTLISKHGVLTQTQATQADSRLEMDVGTLE-----PADTS--TINNLI 494  
 QY 575 SLSSILLANNAVTPNPPNPAQDSHPAIVGS-TTAGSVTISGIPFEDLDDTAYDRYWL 633  
 DB 495 NISSI-----DQAKKAKIETKATSKNLTLSTGTTILLDPTGTFYENHS-IR 538  
 QY 634 SNQKINVLKQL-GTKPPANAPSDLTGLNEMPKYQYQGSW-KLAWDPNTPANNPPTLKAT 691  
 DB 539 NPQSYDILELKASGVTSTAVTPDPIMGKEF-HYGYQGTWGPVWGTGASTTATP-----N 593  
 QY 692 WTKTGYNPQPRVAVSLVPSNLSLWGSTLDIRSAHSAIQASVDGRSVCRLWVSGVSNFFVHD 751

DB 594 WTKTGYNPQPRVAVSLVPSNLSLWNAFIDISSLHYLMETANEGLQGRFACWAGLSNFFHKD 653  
 QY 752 RDALGQGVRYTSGYSGANSYFGS-SMFGLAFTFVGRSKDYVVCRSNHHACIGSVYLS 810  
 DB 654 STKTRRGPRHLSSGGYVIGGNLHTCSDKILSAAFQQLFGDRDRDYFAKNQ-----GTVYGG 708  
 QY 811 TQQAALCGSYLPQDAFI-----PASVQFGNHMKTSYTPABE 846  
 DB 709 TLY-----YQHNETYISLPCKLRPCSLSYVTEIPVLFSGNLSYTHTDNDLTKTYTPT 763  
 QY 847 SDVRWNNCLAGIGAGLPIVITPSKLYNLRLRPFVQAEPSYADHESFTEGDOARAFKS 906  
 DB 764 VKSGWENDSPALEFGGRAPICLDESALF-EQYMPKLOFYVAHQEGFKQGTAREFGS 822  
 QY 907 GHLLNLNVVPGVKFRCSSSTHPNKYSFMAAYICDAYRTISGTETTLTLLSHQBTWTTDAFHL 966  
 DB 823 SRLVNALPIGIRFDKESDCQDATYNTLTGLVTVDLVRSNPDCTTTLRIGSDSKMTFGTNL 882  
 QY 967 ARHGVTVRGSMYASLTNIRVYGHGRVEYRDASRGYGLSAGSRVRF 1012  
 DB 883 ARQALVLRAGNHFCFNSNFAPSQSFELRGSSRNYNVDLGAKYQF 928

RESULT 40  
 ABP56019  
 ID ABP56019 standard; protein; 926 AA.  
 XX  
 AC ABP56019;  
 DT 29-AUG-2003 (revised)  
 DT 25-FEB-2003 (first entry)  
 DE Chlamydia psittaci antigen CP4#12 protein SEQ ID NO:57.  
 KW Chlamydia psittaci; vaccination; vaccine; antigen; immune response;  
 KW immunisation; antibacterial; infection.  
 OS Chlamydophila caviae.  
 XX  
 PN WO200253588-A2.  
 PD 11-JUL-2002.  
 PF 17-DEC-2001; 2001WO-US048715.  
 XX  
 PR 15-DEC-2000; 2000US-00738269.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA Johnston SA, Stemke-Hale K, Sykes KF, Kaltenboeck B;  
 PI  
 XX WPI; 2002-537942/57.  
 DR N-PSDB; ABQ84775.  
 XX  
 PT Vaccine for immunization of animal, preferably bovine, against Chlamydia  
 PT psittaci, comprises at least one polynucleotide having a C. psittaci  
 PT sequence, or at least one C. psittaci antigen.  
 XX  
 PS Claim 9; Page 158-161; 164pp; English.  
 XX  
 CC The present invention describes a vaccine (I) for the immunisation of an  
 CC animal against Chlamydia psittaci comprising at least one polynucleotide  
 CC (Ia) having a C. psittaci sequence, or at least one C. psittaci antigen  
 CC (Ib), and a carrier. (Ia) and (Ib) have antibacterial activity. (I) is  
 CC useful for the immunisation of a bovine. The present sequence represents  
 CC a C. psittaci antigen from the present invention. (Updated on 29-AUG-2003  
 CC to standardise OS field)  
 XX  
 SQ Sequence 926 AA;

Query Match 21.3%; Score 1122; DB 5; Length 926;  
 Best Local Similarity 29.7%; Pred. No. 2.6e-68;  
 Matches 323; Conservative 152; Mismatches 376; Indels 236; Gaps 34;

Qy	1	MOTSFKPFLSMILAYSCSLNGGVAEIMVPQGIYDGETILTVSPPTVI-----	51
Db	1	MRPSLYKLILSTL-----TLPISFHSQLAHEVALTQE	34
Qy	52	-----GDPSTGVTFVSAGELTKLNLDNSI-----AALPLSCFNLGSLFTVLGRHS	97
Db	35	SILDANGAFSPQSTAGG-TIYNVEDSIVDVGQTAALASSAFVQTADNLTFKGNHS	93
Qy	98	LTFENIRTSNGAALSNAADGLFTIEGKELSFNCNLLAVLPAATYKGSQTPPTS	157
Db	94	LSITNANAGANPAGINVTADKILTLTDFSKLSFKEC-----PSSLVNTG-----	138
Qy	158	TFSNGTIYSKTDLLILNNEKFSFYNLVSOGDAIDAKSLTVQGISKLCVFOENTAQADG	217
Db	139	-----KGAKSGGALNLANNASILFDQYSAENGGAISCKAFSLTGSKEISFTTNSTAKG	195
Qy	218	GACQVVTFSAMANEAPIAFVANVAGVGGIAAVODGQGVSSSTSTEDPVVPSRNTA	277
Db	196	GA-----IAATGIAHLSDNQ-----GTIRFSGNTA	220
Qy	278	VFPDGNVAVGGIYSGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGOASNTSNYDGG	337
Db	221	V-----NSGGAVYSEASMTIAGNNHVAFSNNAVS-----GSSDGGC	256
Qy	338	GAIFCKNGAQAAGSNNGSVSPDGEVVPFSSNVAAAGGGAIIYAKKLSVANCGPVQFLRN-	396
Db	257	GAHCSK-----TGSAPTILTRDNKVLIPENTSSAKGGAITYDKLILTSGGGTAFINNK	311
Qy	397	---IANDGGAIVLGBESGELSADYDGIIPDGNLKRKTAKENAADVNGVTVSSQAIASMSG	453
Db	312	VTHATPKGAIGIANGEGCSLTAHGDITFDNNLMAT-QDNA-----TIKRNAINIEGN	364
Qy	454	GKITTILRAKAGHOILFNDPIEMANGNNQPAQSKLLKINDGEG---YTGDIVPA-----	504
Db	365	GKPVNLRASAGKTI SFVDPIV-EGN-----AADLLTNKAEGDKTYNGRIIPSGEKLTE	418
Qy	505	-----NGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQSGSL-YMEAGSTWDFVTPQP	556
Db	419	EQAADVADNLKTTPTQPIITLAAGELVLRSGVEBAKTVVQFAGSLILMDAGT-----	469
Qy	557	PQPPPAANQLITLSNLHLSLSLANNVATNPPTNPPAODSHP-AVIGSTTAG-SVTISG	614
Db	470	-----KLSAKTEATLTNLAINENTLDGKKFPAVVDAAAGKNVTLSC	511
Qy	615	PIFFEDLDDTAYDRYDWLGSNQKIN-----VLKLQL-----GTKPPANAPSDLTGNEMPKYG	667
Db	512	AIGVIDPTGKFYE-----NHKLNDTLALGGIQLSGKGSVTTTNPVPSHV-VGVAETHYG	563
Qy	668	YQGSWKLAWDPNTANNGPYTLKA--TWTKTYNPGPERVASLYPNLSLWGSILDIRSAHSA	725
Db	564	YQGNWSVSWYKDN-NSDPKTQTAFITWNKTGYVPNPERRAPLVNLSLWGSFIDLRISQDV	622
Qy	726	IQASVDG-RSYCGLWGSVSNPFYHDDRDALGQYRYSIGYSILGANSYFG-SSMFGIAP	783
Db	623	LERSVDSILETRGLWVSGIGNPFHKORNAENKFRHISSGYVLGATNTSREDSLSVAP	682
Qy	784	TEVFGRSKDYVCRSNHHACIGSVY-----LSTQQALCGSY-----LFGDAF	825
Db	683	COLPAKDKDYLVSKAANVYAGSVYQYHVSKEFDDTLFLNGPNTCCSGPSKEIPIPLDAQ	742
Qy	826	IRASYGFGNOHMTSYTFABESDVWRDNNCLAGEIGAGLPVITTPSKLYLNEILPFPVQAE	885
Db	743	I--TYCHTANNMTTSYTDYEVKSGWGNDTLGLTLSTSVPIPVFSSSIP-DSYAPFAKLQ	799
Qy	886	PSYADHESPTTEGQARAFKSGHLNLSVPVGVKFDRCSSTHPNKYSFMAAYICDAYRTI	945
Db	800	VVYAHQDDFKBPTTEGRVFESSDLLNSVPIGIKFEKLSYGERSAIDLTLMYIPDVYRHN	859
Qy	946	SGTETLLSHOETWTTFAPLHARGVVRGSMYASLTSNIEVYGHRYEYRDASRGYGLS	1005
Db	860	PSCMTGLAINDVSWLTATNLARQAFIVRAGNHIALTSVEMFSQFGFELRSSSRNYNVD	919

Qy 1006 AGSRVRF 1012

Db 920 LGAKVAF 926

Search completed: May 13, 2006, 12:15:00

Job time : 200 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:17:59 ; Search time 47 Seconds  
(without alignments)  
2071.731 Million cell updates/sec

Title: US-10-701-844-2  
Perfect score: 5267  
Sequence: 1 MQTSFHKFPLSMILAYSCS.....YEURDARGYGLSAGSRVRF 1012

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*  
1: Pir1:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5131.5	97.4	1013	2 G71460	probable outer mem
2	3845	73.0	987	2 H81722	polymorphic membra
3	1365.5	25.9	973	2 B86547	polymorphic outer
4	1365.5	25.9	973	2 F72076	polymorphic outer
5	1365.5	25.9	935	2 C81593	polymorphic membra
6	1101	20.9	928	2 G86546	polymorphic outer
7	1101	20.9	928	2 G81591	polymorphic membra
8	1090	20.7	928	2 D72077	polymorphic outer
9	1090	20.7	928	2 H86546	polymorphic outer
10	1090	20.7	949	2 F81591	polymorphic membra
11	1083	20.6	930	2 D86546	polymorphic outer
12	1083	20.6	930	2 A81591	polymorphic membra
13	1079	20.5	930	2 D72078	polymorphic outer
14	1041	19.8	936	2 C72078	polymorphic membra
15	1040	19.7	936	2 B81591	polymorphic outer
16	1040	19.7	936	2 C86546	polymorphic membra
17	1021	19.4	928	2 B72077	polymorphic outer
18	1021	19.4	928	2 B86546	polymorphic outer
19	1012	19.2	1276	2 B86546	polymorphic membra
20	1012	19.2	1276	2 C81591	polymorphic outer
21	985.5	18.7	1407	2 B72078	polymorphic outer
22	978.5	18.6	772	2 H86492	Pmp_3 [imported] -
23	967	18.4	922	2 B72131	polymorphic outer
24	967	18.4	922	2 B86491	polymorphic membra
25	966	18.3	922	2 F81539	polymorphic membra
26	939.5	17.8	841	2 B72130	polymorphic membra
27	789	15.0	867	2 F81721	polymorphic membra
28	771.5	14.6	878	2 B71460	probable outer mem
29	703	13.3	712	2 B86492	polymorphic outer

## ALIGNMENTS

## RESULT 1

G71460  
probable outer membrane protein G - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
C:Accession: G71460  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1999  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra A:Reference number: A71570; MUID:99000809; PMID:9784136  
A:Accession: G71460  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1013 <ARN>  
A:Cross-references: UNIPROT:O84879; UNIPARC:UPI0000131CF5; GB:AE001360; GB:AE001273; NIT A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: pmpG

Query Match	97.4%;	Score	5131.5;	DB 2;	Length	1013;			
Best Local Similarity	97.0%;	Pred. No.	8.6e-304;						
Matches	983;	Conservative	18;	Mismatches	11;	Indels	1;	Gaps	1;
<hr/>									
QY	1	MQTSFHKFPLSMILAYSCS	INGGGYAAEIMVPOGIYDGETLTVSPYTVIGDPGSGTVTF	60					
DB	1	MQTSFHKFPLSMILAYSCS	LSGGYAAEIMIPQGIYDGETLTVSPYTVIGDPGSGTVTF	60					
<hr/>									
QY	61	SAGELTLKNLNDNSIAALPLSCF	GNLLGSFTVLGRGHSLTTFENIRTSNCAALSNSAAGL	120					
DB	61	SAGELTLKNLNDNSIAALPLSCF	GNLLGSFTVLGRGHSLTTFENIRTSNCAALSNSAAGL	120					
<hr/>									
QY	121	FTIEGPKELSFNCNSLLAVLPAATNKGSGQPTTSTPSNGTIYSKTDLLILNNEKPSF	180						
DB	121	FTIEGPKELSFNCNSLLAVLPAATNKGSGQPTTSTPSNGTIYSKTDLLILNNEKPSF	180						
<hr/>									
QY	181	YSLNLSGDDGGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTFSAMANEAPIAFVAN	240						
DB	181	YSLNLSGDDGGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTFSAMANEAPIAFVAN	240						
<hr/>									
QY	241	VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVSFRNTAVFDGNVARVGGGIYSYGNVAF	300						
DB	241	VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVSFRNTAVFDGNVARVGGGIYSYGNVAF	300						
<hr/>									
QY	301	NGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYGGGAIFCKNGAQ-AGSNNSGVSFD	359						
DB	301	NGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAAGSNNSGVSFD	360						
<hr/>									
QY	360	GEQVWFSSNVAAGKGGAIYAKKLSVANGCPVOLFNIANDGGAIVLGSSELGSADYG	419						
DB	361	GEQVWFSSNVAAGKGGAIYAKKLSVANGCPVOLFNIANDGGAIVLGSSELGSADYG	420						
<hr/>									
QY	420	DIIFDGNLKRKATENAADVGVTVSSQAISMGSGGKITTLRAKAGHQITLFDNPITEMANGN	479						

Db	421	DIIIFDGNLXRTAKENAAADVNGVTYSSOAISMSGGKITTFRAGAGHOILFNDPIEMANGN	480
Qy	480	NPAQSCKLLKINDGEGVTGDIVPANGSSTLYQNVTIEQGRIVLREKAKUSVNSLSOTGG	539
Db	481	NPAQSSEPLKINDGEGVTGDIVPANGNSTLYQNVTIEQGRIVLREKAKUSVNSLSOTGG	540
Qy	540	SLYNHAGSTWDFVTVPQPQPAPAAQLITLSNLHLSSLIANNAVNTPPTPPAQDSHP	599
Db	541	SLYNHAGSTLDFVTVPQPQPAPAAQLITLSNLHLSSLIANNAVNTPPTPPAQDSHP	600
Qy	600	AVIGSTTAGSVTISGPPIFFEDDLDTAYDRYDWLGSNOKINVLKLGLTKPKPANAPSDDLTL	659
Db	601	AIIGSTTAGSVTISGPPIFFEDDLDTAYDRYDWLGSNOKIDVLKLGLGTQPSANAPSDLTL	660
Qy	660	GNEPMKPYGQGSWKLANDPNTANNPGPYTLKATWTKTGYNCPGERVASLVNSLMGSLTLDI	719
Db	661	GNEPMKPYGQGSWKLANDPNTANNPGPYTLKATWTKTGYNCPGERVASLVNSLMGSLTLDI	720
Qy	720	RSAHSATQASVDGRSYCRGLMWGVSNPFYHDRDALGQGYRISGGYSLGANSYFGSGSMF	779
Db	721	RSAHSATQASVDGRSYCRGLMWGVSNPFYHDRDALGQGYRISGGYSLGANSYFGSGSMF	780
Qy	780	GLAFTEVFGRSKOYVVCRSNHHCIGSVYSTQALCGSYLFGDAFTRASYGFGNQHMKT	839
Db	781	GLAFTEVFGRSKOYVVCRSNHHCIGSVYSTQALCGSYLFGDAFTRASYGFGNQHMKT	840
Qy	840	SYTPAESDVRMWNCLAGHIGAGLPVITPSKLYLNELRPFFVOAERFSYADHESFTBECD	899
Db	841	SYTPAESDVRMWNCLAGHIGAGLPVITPSKLYLNELRPFFVOAERFSYADHESFTBECD	900
Qy	900	QARAFKSHLINLSVPVGKFDRCSTHPNKYSFMAAYICDAYRTISGTETILLSHOETW	959
Db	901	QARAFKSHLINLSVPVGKFDRCSTHPNKYSFMAAYICDAYRTISGTETILLSHOETW	960
Qy	960	TTFDAFLRHGVIIRGSGMYASLTNIENVYGHGRYEYRDASRGYGLSAGSRVRF	1012
Db	961	TTFDAFLRHGVIIRGSGMYASLTNIENVYGHGRYEYRDASRGYGLSAGSRVRF	1013

RESULT 2

H81722

polymorphic membrane protein G family TC0263 [imported] - Chlamydia muridarum (strain Nijmegen)

C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004

C;Accession: H81722

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.L.; et al. Nucleic Acids Res. 28, 1397-1406, 2000

A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A;Reference number: A81500; UID:20150255; PMID:10684935

A;Accession: H81722

A>Status: preliminary

A:Molecule type: DNA

A;Residues: 1-987 <TET>

A;Cross-references: UNIPROT:Q9PL45; UNIPARC:UIP0000057867; GB:AEO02293; GB:AEO02160; NID:1000000000

A;Experimental source: strain Nigg (MoPn)

C;Genetics:

A;Gene: TC0263

Query Match		73.0%; Score 3845; DB 2; Length 987;
Best Local Similarity		72.6%; Pred. No. 1.3e-225;
Matches	739; Conservative	105; Mismatches 136; Indels 38; Gaps 13;

  

Qy	1	MQTSPFHKKFLLSMILAYSCCSLNGGYYAAEIMVPGIYDGFTLTVSPPTVIGDPGPTVTF	60
Db	2	MQTSPFHKKFLLSMILAYSCCSLNGGYYAAEIMVPGIYDGFTLTVSPPTVIGDPGPTVTF	58
Qy	61	SAGELTILKNLONSIAALPLSCPGLNLGSPFTVLGRGHSLTPENTRTSTNGALSNSAADGL	120
Db	59	SSGSELKUNLONSIAALPLSCPGLNLGSPFTVLGRGHSLTPENTRTSTNGALSNSAADGL	118
Qy	121	FTTEGFKELFSPNCNSLIALLVLAATTNKGSGT-PTTTSTPSNGTIYSKTDILLINNKFPS	179

A:Residues: 1-973 <STO>  
A:Cross-references: UNIPROT:Q9Z896; UNIPARC:UPI000004707B; GB:BA000008; NID:98978824; PI  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: pmp\_13  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 25.9%; Score 1365.5; DB 2; Length 973;  
Best Local Similarity 33.5%; Pred. No. 5.8e-75;  
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY 1 MOTSFHKKPLSLMILAYSCCSLNGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP 60  
DB 1 MTSIRKPLISTTLAPCPAST---AFTVEIMPSNFDDSGSKI-PPYTTLSDPRTGLCI 56

QY 61 SAGELTNLKNDLSIAALPLSCFNLGSLFTVLRGHSHTFENIRSTNGAALSNAADG- 119  
DB 57 FSGDLVIANLDAISRTSSCFNRAGALQILKGGVFLNIRSSADGAIAISVITQNP 116

QY 120 ---LFTIEGKELSFNCNLSLAVLPAAATNKGSTPTTSTPSNGTIYSKTDLLLNNE 176  
DB 117 ELCPLSFSQSFQIMFNCESLT-----SDTSASNVIPIHSAIYATTPMLPTND 165

QY 177 KFSFYSLVSGDGAIDAKSLTVQGISKLCVQFQNTAQAADGGACQVVTFSAMANEAPIA 236  
DB 166 SILFOYRRSAGFGAALRGTSITIENTKSLFNGNGSISNGGALTGSAALINLNSAPVI 225

QY 237 FVANVAGVGGGTAIVQDQGGVSSSTSTEDPVVFSRNTAVEFDGCVNARVGGIYSYN 296  
DB 226 FSTNATGIYGGAIYLT-----GGSMILTS-----GNLSGVLFVNNSSRSGAIYANGN 272

QY 297 VAPLNGKTLFLNNVASPV-YIAAKOPTSQASNTSNYGDGGAIFCKNGAQSNNSSG 355  
DB 273 VTFNNSDLTTFQNTASPNLSPPATPPPTPAVTP-LLGYGGAICTPPTPPGCV-S 330

QY 356 VSPDGGVGFSSNVAAGKGAIYAKKLSVANCGPVQFLNIAVDGGAIVLGSSELSS 415  
DB 331 LTISSGNSVTFLENIASEQGALYKKISIDSNKSTIFLNGTAGKGAIAIPESGELS 390

QY 416 ADYGDIIIFDGNLKRKTAKENAAVNGVTVSSQAISSMSGGKITTLRAKAGHQILFNDPIEM 475  
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGAATQGYTLFYDPI-- 439

QY 476 ANGNQPAQSSKLLKIN-----DREGYTGDIVP-----ANGSSTLYQNVTIE 517  
DB 440 TSDDLGAASAAATVWVNPKASADG-AYSGTIVFSGETLTATEAATPANATSTLNQKLE 498

QY 518 QGRIVLRKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQPPQPPAANQILITLSHL 576  
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGAATQGYTLFYDPI-- 439

QY 476 ANGNQPAQSSKLLKIN-----DREGYTGDIVP-----ANGSSTLYQNVTIE 517  
DB 440 TSDDLGAASAAATVWVNPKASADG-AYSGTIVFSGETLTATEAATPANATSTLNQKLE 498

QY 518 QGRIVLRKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQPPQPPAANQILITLSHL 576  
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGAATQGYTLFYDPI-- 439

Db 830 SSRILNLKQIIPFVKAEVAYATHGGIQENTPEGRIFGHGHLNVAVPVGVFRGNHNR 889  
QY 929 NKYSFMAAYICDARTTISGTETILLSHOETWTTFADFLHARGVVRGSMYASLTNIEVY 988  
Db 890 DFYTIIVAYADVTNRNPDCTTLPINGATWTSIGNNLTRSTLLVQASHTSVNDVLEIF 949  
QY 989 GHGRIEYRDASRGYGLSAGSRVP 1012  
Db 950 GHGCDIRRTSKQVTLDIGSKLRF 973

RESULT 4

F72076  
polymorphic outer membrane protein g family - Chlamydomophila pneumoniae (strain CWL029)  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: F72076  
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.,  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: F72076  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-973 <ARN>  
A:Cross-references: UNIPROT:Q9Z896; UNIPARC:UPI000004707B; GB:AE001629; GB:AE001363; NTI  
A:Experimental source: strain CWL029  
C:Genetics:  
A:Gene: pmp\_13  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 25.9%; Score 1365.5; DB 2; Length 973;  
Best Local Similarity 33.5%; Pred. No. 5.8e-75;  
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY 1 MOTSFHKKPLSLMILAYSCCSLNGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP 60  
DB 1 MTSIRKPLISTTLAPCPAST---AFTVEIMPSNFDDSGSKI-PPYTTLSDPRTGLCI 56

QY 61 SAGELTNLKNDLSIAALPLSCFNLGSLFTVLRGHSHTFENIRSTNGAALSNAADG- 119  
DB 57 FSGDLVIANLDAISRTSSCFNRAGALQILKGGVFLNIRSSADGAIAISVITQNP 116

QY 120 ---LFTIEGKELSFNCNLSLAVLPAAATNKGSTPTTSTPSNGTIYSKTDLLLNNE 176  
DB 117 ELCPLSFSQSFQIMFNCESLT-----SDTSASNVIPIHSAIYATTPMLPTND 165

QY 177 KFSFYSLVSGDGAIDAKSLTVQGISKLCVQFQNTAQAADGGACQVVTFSAMANEAPIA 236  
DB 166 SILFOYRRSAGFGAALRGTSITIENTKSLFNGNGSISNGGALTGSAALINLNSAPVI 225

QY 237 FVANVAGVGGGTAIVQDQGGVSSSTSTEDPVVFSRNTAVEFDGCVNARVGGIYSYN 296  
DB 226 FSTNATGIYGGAIYLT-----GGSMILTS-----GNLSGVLFVNNSSRSGAIYANGN 272

QY 297 VAPLNGKTLFLNNVASPV-YIAAKOPTSQASNTSNYGDGGAIFCKNGAQSNNSSG 355  
DB 273 VTFNNSDLTTFQNTASPNLSPPATPPPTPAVTP-LLGYGGAICTPPTPPGCV-S 330

QY 356 VSPDGGVGFSSNVAAGKGAIYAKKLSVANCGPVQFLNIAVDGGAIVLGSSELSS 415  
DB 331 LTISSGNSVTFLENIASEQGALYKKISIDSNKSTIFLNGTAGKGAIAIPESGELS 390

QY 416 ADYGDIIIFDGNLKRKTAKENAAVNGVTVSSQAISSMSGGKITTLRAKAGHQILFNDPIEM 475  
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGAATQGYTLFYDPI-- 439

QY 476 ANGNQPAQSSKLLKIN-----DREGYTGDIVP-----ANGSSTLYQNVTIE 517  
DB 440 TSDDLGAASAAATVWVNPKASADG-AYSGTIVFSGETLTATEAATPANATSTLNQKLE 498

QY 518 QGRIVLRKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQPPQPPAANQILITLSHL 576  
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGAATQGYTLFYDPI-- 439



Db 499 GGTALRNAGATLVNHFQDEKSVVIMDAGTT--LATNGANNITDGA---ITLNLKLVINL 553  
Qy 577 SLLANNVTPPTNPPAQDSHPAIVG-STTAGSVTISGPIFFEDLDATAYDRYDMLGSN 635  
Db 554 DSLDGTKA-----AVVNVQSTNGALTISGTLGLVKNQSCDCHGMFNKD 598  
Qy 636 -QKINVLKQLGTLKPPANAPSDL-TLGNEMPKYGYGGSWKLAWDPNTANNPYTLKATWT 693  
Db 599 LQQVPILELKATSNVTITDPSLGTNGYQGSYGYGTWETIDTT-----HTVTGNWK 653  
Qy 694 KTYNPGPVERVASLVPNSLWGSILDIRSAHSAIOASVDGRSY-CRGLMVSGVSNFFYHNR 752  
Db 654 KTYLPHPERLAPLIPNSLMANVIDLRAVSA--SAADGEDVPGKQLSITGITNFFHANH 711  
Qy 753 DALQGGYRYISGGYSLGANSY---FGSMPGLAFTEVFGRSKDYVVCNHHACITGSVYL 809  
Db 712 TGDARSYRHMGGYLI--NTYTRITPDAALSGLPGQLFTKSKDYLVGHGSHNVYATVYS 769  
Qy 810 STQOALCG-SYLPFGDAFIRASVYFGNOMKTSYTFABESDVVRWNNCLAGEIGLPIVI 868  
Db 770 NITKSLFGSSRRFFSGTGRVTSRNEKVKTSYTKLPKGRCSWSNCCWLGEGLNLPITL 829  
Qy 869 TPSKLYLNELRPFVQAFPSYADHESFTEEGDQARAFKSHLNLSPVGVKFDRCSSHP 928  
Db 830 SRRILNLKQLIPPVKARAYATHGGIOENTPEGRIFGCHLLNVAVPVGVFRGKSHNRP 889  
Qy 929 NKYSFMAAYICDAYRTISGTTETLLSHOETWTDDAPHLARHGTVVRGSMYASLTSNIEVY 988  
Db 890 DFTYIIIVAYADPVYRHNPDCTTLPLNGATWTSIGNNLTRSTLLVQASSTSVNDVLEIF 949  
Qy 989 GHGREYRDASRGYGLSAGSRVRF 1012  
Db 950 GHGCCDIRTSRQYTLDIGSKLRF 973

RESULT 5  
C81593  
polymorphic membrane protein G family CP0299 [imported] - Chlamydia pneumoniae (stra  
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 15-Jun-2001  
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, M.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A/Reference number: A81500; MUID:20150255; PMID:10684935  
A/Accession: C81593  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-995 <REA>  
A/Cross-references: UNIPARC:UPI00001655F9; GB:AE002191; NID:g7189216; PIDN:  
A/Experimental source: strain AR39, HL cells  
C/Genetics:  
C/Genes: CP0299  
C/Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Query Match 25.9%; Score 1365.5; DB 2; Length 995;  
Best Local Similarity 33.5%; Pred. No. 6e-75;  
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

Qy 1 MCTSPHKFPLSMILAYSCSLNGGYAAEIVVQGIYDGEILTVSPVTVIGDPSGTTVP 60  
Db 23 MKTSIRKFLISTLAPCPAST---AFTVEVIMPSNFDSGSKI--FPYTTLSDPRGTLCI 78  
Qy 61 SAGELTLKLNLSIALPLSCFNLGSGFTVLGRGHSITFENIRTSWTGAALSNSAAG- 119  
Db 79 FSGDLYIANLDAISRTSSCSFNSRAGALQILKRGVGFNIRISADGAALISSVITQNP 138  
Qy 120 ---LFTIBGFKELSFNSCNLSLAVLPAATAATNKGSTPTTTTPSPNGTTYSKTDLLLANE 176  
Db 139 ELCPLSPSGFGFMIDNCESLT-----SDTSASNVIPHASAIYATTPMLFTNND 187  
Qy 177 KFSFYSNLVSGDGGDAIDAKSLTVQGISKLCVFOENTAQADGACQVTSFSMANEAPIA 236

Db 188 SILFOYNRSAFGAAIRGTSITIENTKKSLFNGNGSISNGGALTGSAAINLNNASVI 247  
Qy 237 FVANVAGVRGGGIAAODGQGVSSSTSTEDPVVSPFRNTAVBFDGNVARVGGGIYSYCN 296  
Db 248 FSTNATGIYGAAYLT-----GGSMLTS-----GNLSGVLFVNNSSRGGAIYANGN 294  
Qy 297 VAFNLNGKTLFLNNAVAPV-YIAAKQPTSGQASNTSNNGYGGGAIFCRKGAAQASNNNGS 355  
Db 295 VTFNNSDLTFQNTATSPQNSLPAPTPTPPAATP-LLGYGGAIFCTPTPATPTPTGTV-S 352  
Qy 356 VSPGEGVVPFSSNVAACKGAIYAKLVANCPVQVPLRNIANDGGAIYLGESGELSLS 415  
Db 353 LTISSGNSVTPLENIASEGGALYKKISIDSNKSTIPLGNTAGKGAIAIPESGELSLS 412  
Qy 416 ADYGDIIIPDKMLKKTAKENAADVNGVTVSSQAIISWGSQGIITLRKAKAGHOILFNDPIEM 475  
Db 413 ANQGDILLFNKLSITSG-----TFTRNSIHFKDAKFAITLGCATGYTLYFYDPI-- 461  
Qy 476 ANGNQPAQSSKLLKIN-----DGEYTGDIIVF-----ANGSSTLYQNVITR 517  
Db 462 TSDDLSSAASAAATVVPKASADG-AYSGTIVFSGETILTATEAATPANATSTLNQKLE 520  
Qy 518 QGRIVLREKAKLSYNSLSQSGS-LYMEAGSTWDPVTPPQPPQPPAANQLITLSHLISL 576  
Db 521 GGTALRNAGATLVNHFQDEKSVVIMDAGTT--LATNGANNITDGA---ITLNLKLVINL 575  
Qy 577 SLLANNVTPPTNPPAQDSHPAIVG-STTAGSVTISGPIFFEDLDATAYDRYDMLGSN 635  
Db 576 DSLDGTKA-----AVVNVQSTNGALTISGTLGLVKNQSCDCHGMFNKD 620  
Qy 636 -QKINVLKQLGTLKPPANAPSDL-TLGNEMPKYGYGGSWKLAWDPNTANNPYTLKATWT 693  
Db 621 LQQVPILELKATSNVTITDPSLGTNGYQGSYGYGTWETIDTT-----HTVTGNWK 675  
Qy 694 KTYNPGPVERVASLVPNSLWGSILDIRSAHSAIOASVDGRSY-CRGLMVSGVSNFFYHNR 752  
Db 676 KTYLPHPERLAPLIPNSLMANVIDLRAVSA--SAADGEDVPGKQLSITGITNFFHANH 733  
Qy 753 DALQGGYRYISGGYSLGANSY---FGSMPGLAFTEVFGRSKDYVVCNHHACITGSVYL 809  
Db 734 TGDARSYRHMGGYLI--NTYTRITPDAALSGLPGQLFTKSKDYLVGHGSHNVYATVYS 791  
Qy 810 STQOALCG-SYLPFGDAFIRASVYFGNOMKTSYTFABESDVVRWNNCLAGEIGLPIVI 868  
Db 792 NITKSLFGSSRRFFSGTGRVTSRNEKVKTSYTKLPKGRCSWSNCCWLGEGLNLPITL 851  
Qy 869 TPSKLYLNELRPFVQAFPSYADHESFTEEGDQARAFKSHLNLSPVGVKFDRCSSHP 928  
Db 852 SRRILNLKQLIPPVKARAYATHGGIOENTPEGRIFGCHLLNVAVPVGVFRGKSHNRP 911  
Qy 929 NKYSFMAAYICDAYRTISGTTETLLSHOETWTDDAPHLARHGTVVRGSMYASLTSNIEVY 988  
Db 912 DFTYIIIVAYADPVYRHNPDCTTLPLNGATWTSIGNNLTRSTLLVQASSTSVNDVLEIF 971  
Qy 989 GHGREYRDASRGYGLSAGSRVRF 1012  
Db 972 GHGCCDIRTSRQYTLDIGSKLRF 995

RESULT 6  
G86546  
polymorphic outer membrane protein G family [imported] - Chlamydia pneumoniae (stra  
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: G86546  
R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I  
Nucleic Acids Res. 28, 2311-2314, 2000  
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A/Reference number: A86491; MUID:20330349; PMID:10871362  
A/Accession: G86546  
A/Status: preliminary  
A/Molecule type: DNA

A:Residues: 1-928 <STO>  
A:Cross-references: UNIPROT:O9RB65; UNIPARC:UPI000002FFP0; GB:BA000008; NID:g8978821; PI  
A:Experimental source: strain J138  
C:Genetics:  
C:Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Query Match 20.9%; Score 1101; DB 2; Length 928;  
Best Local Similarity 30.5%; Pred. No. 6.4e-59;  
Matches 324; Conservative 150; Mismatches 402; Indels 188; Gaps 34;

QY 1 MOTSPHKFPLSMILA-YSCCSLNGGGYAAEIMVPGIYDGETLVSPFYTVIGDPSGTVV 59  
DB 1 MKSQFSLVLSSTLACFTSCSTVFAA-TAENIGPSDSFDGNTGT--YTPKNTTGTIDY 57

QY 60 PSAGBLTKNLDSNLAALPLSCFNLGSPVVLGRHSLTPENIRTSNGALSNSAAG 119  
DB 58 TLTGDTITLQNLGDS-AALTKGCFSDTTESLSPAGKGYSLFLNKSABGAALS-VTTDK 115

QY 120 LFTIEGKELSPSNCNLSLAVLPAATTNKGSTPTTTSPSNGTIYSKTDLLLNNEKFS 179  
DB 116 NLSLTFGSSLT--LAAPSVITTPSG-----KGAVKCGDLTFDNNGTIL 159

QY 180 FYSNLVSGDGGDAIDAKSLTVQGISKLCVQFQNTAQDGGACQVWTSFSAMANEAPIAFA 239  
DB 160 FKQDYCEENGGAISTKNLSLKNSTGSIPEGNKSSA----- 195

QY 240 NVAGVRGGIAAVQDQGGVSSSTSTEDPVVFSRNTA-VEFDGNVAR-VGGGIYSYGNV 297  
DB 196 --TGKKGGAICA-----TGT-----VDITNTAPTFLFSNNIAEAGGAINSTGNC 238

QY 298 AFLNNGKTLPLNVAAPVYIAAKQPTSGQASNTSNNGDGGALFCNGAQAQSNNSGSVS 357  
DB 239 TITGNTSLVFSN-----SVTATAGNGGAL-----SGDADVT 270

QY 358 FDGEGVFPSSNVAAGKGAIIYAKKLISVAN--CGPVQFLRNI-----ANDGGAIIYLGSG 410  
DB 271 ISGNQSVTFSGNQAVANGAIIYAKKLTLASGGGGISFSNNIVQGTAGNGGAISILAAG 330

QY 411 ELSLSADYGDIIIPDGN-----LKRTAKENADVNGVTVSSQAISMGSQGIITLRAKAGH 465  
DB 331 ECSSLAEAGDITFNGNAIVATTPTQTKRNSIDI-----GSTAKITNLRAISGH 378

QY 466 QILFNDPIEMANGNNOPOASSKLLKINDGEG-----YTGDIVPA-----NGS 507  
DB 379 SIFFYDPITA-----NTAADSTDTLNLNKADAGNSTDYSGISVFSGEKLSDEAKVADNLT 434

QY 508 STLYQNVTIEQGRIVLREKAKLSVNSLSQT--GGSLYMEAGSTWDFVTPQPPQPPAANQL 566  
DB 435 STLKQPVTLTAGNLVLRGVTLTKGFTQTAGSSVIMDAGTTL-----KASTE 483

QY 567 ITLSNLHLSSLILLANNVTPNPPAQSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626  
DB 484 VTILGLSIPVDSL-----GEGKKVYIAASAASAKNVALSGPILLIDNQNAY 529

QY 627 DRYDMLGSKNQKLVNKLQ-LGTHKPPANASDLTLGNEMPKYGYQGSWKLAWDPTANNCP 685  
DB 530 ENHD-LGKTQDPSFVQLSALGATATTTDVAAPTVAATP-THYGYQGTWGMVDDDTAST-P 586

QY 686 YTLKAT--WTKGYNPQPERVASLVNLSLWGSILDIRSAHSAIQASVDGRSVCRLGWVG 743  
DB 587 KTKTATLWNTGYLNPENRQGPLVPSNLMSGFSDIQAIGVTERALTLCSDRGFWAAG 646

QY 744 VSNFFYHRRDALQGGYRISGGYSL--GANSYFGSSMPGLAFTRVPGRSKYDYVCRSNHHA 802  
DB 647 VANFLDKDKGKRYKHKSGGYAIGGAAQTCSENILISPAFCQLFGSKDPLVAKNHTDT 706

QY 803 CIGSVYLS-----TQOALCGSYLPEDAFI--RASYGFQNGHMTSYFAESDV 849  
DB 707 YAGAFYIQTHTGCSFGTICLLDLKPLGWSHKLPLVLEQLAYSHSVNDLTKTYATPEVK 766

QY 850 RWNWNLACIGAGLPIVITPSKLY-LNELRPPVQAEFSYADHESFTTEEGDQARAFKSGH 908

DB 767 SWGNNAFMMLGASGSHY--PEYLHCFDPTVAPYIKLNLTYIRQDSFSEKTEGSRFDDSN 824

QY 909 LLNLSPVGVKDFRCSSSTHPNKYGFMAAYICDAYRTISGTETTLSSHQETWTTDAFHAR 968  
DB 825 LPNLSLPIGVKFEKSCDNCPSYDLTSLSYVFDLIRNDPKCTALTALVISGASWETVANNLAR 884

QY 969 HGVVVRGSMYASLTNSIEVYGHGRYERDASRGGLSAGSRVPF 1012  
DB 885 QALQVRAGSHYAFSPMEFVLQGFVEVRGSSRIYNVDLGGKPFQF 928

RESULT 7  
GB1591  
polymorphic membrane protein G family CP0303 [imported] - Chlamydia pneumoniae (stra  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: GB1591  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey,  
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: AB1500; MUID:20150255; PMID:10684935  
A:Accession: GB1591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-928 <REA>  
A:Cross-references: UNIPROT:O9RB65; UNIPARC:UPI000002FFP0; GB:AE002161; NIT  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
C:Gene: CP0303  
C:Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Query Match 20.9%; Score 1101; DB 2; Length 928;  
Best Local Similarity 30.5%; Pred. No. 6.4e-59;  
Matches 324; Conservative 150; Mismatches 402; Indels 188; Gaps 34;

QY 1 MOTSPHKFPLSMILA-YSCCSLNGGGYAAEIMVPGIYDGETLVSPFYTVIGDPSGTVV 59  
DB 1 MKSQFSLVLSSTLACFTSCSTVFAA-TAENIGPSDSFDGNTGT--YTPKNTTGTIDY 57

QY 60 PSAGBLTKNLDSNLAALPLSCFNLGSPVVLGRHSLTPENIRTSNGALSNSAAG 119  
DB 58 TLTGDTITLQNLGDS-AALTKGCFSDTTESLSPAGKGYSLFLNKSABGAALS-VTTDK 115

QY 120 LFTIEGKELSPSNCNLSLAVLPAATTNKGSTPTTTSPSNGTIYSKTDLLLNNEKFS 179  
DB 116 NLSLTFGSSLT--LAAPSVITTPSG-----KGAVKCGDLTFDNNGTIL 159

QY 180 FYSNLVSGDGGDAIDAKSLTVQGISKLCVQFQNTAQDGGACQVWTSFSAMANEAPIAFA 239  
DB 160 FKQDYCEENGGAISTKNLSLKNSTGSIPEGNKSSA----- 195

QY 240 NVAGVRGGIAAVQDQGGVSSSTSTEDPVVFSRNTA-VEFDGNVAR-VGGGIYSYGNV 297  
DB 196 --TGKKGGAICA-----TGT-----VDITNTAPTFLFSNNIAEAGGAINSTGNC 238

QY 298 AFLNNGKTLPLNVAAPVYIAAKQPTSGQASNTSNNGDGGALFCNGAQAQSNNSGSVS 357  
DB 239 TITGNTSLVFSN-----SVTATAGNGGAL-----SGDADVT 270

QY 358 FDGEGVFPSSNVAAGKGAIIYAKKLISVAN--CGPVQFLRNI-----ANDGGAIIYLGSG 410  
DB 271 ISGNQSVTFSGNQAVANGAIIYAKKLTLASGGGGISFSNNIVQGTAGNGGAISILAAG 330

QY 411 ELSLSADYGDIIIPDGN-----LKRTAKENADVNGVTVSSQAISMGSQGIITLRAKAGH 465  
DB 331 ECSSLAEAGDITFNGNAIVATTPTQTKRNSIDI-----GSTAKITNLRAISGH 378

QY 466 QILFNDPIEMANGNNOPOASSKLLKINDGEG-----YTGDIVPA-----NGS 507  
DB 379 SIFFYDPITA-----NTAADSTDTLNLNKADAGNSTDYSGISVFSGEKLSDEAKVADNLT 434

QY 508 STLYQNVTIEQGRIVLREKAKLSVNSLSQT--GGSLYMEAGSTWDFVTPQPPQPPAANQL 566  
DB 435 STLKQPVTLTAGNLVLRGVTLTKGFTQTAGSSVIMDAGTTL-----KASTE 483

QY 567 ITLSNLHLSSLILLANNVTPNPPAQSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626  
DB 484 VTILGLSIPVDSL-----GEGKKVYIAASAASAKNVALSGPILLIDNQNAY 529

QY 627 DRYDMLGSKNQKLVNKLQ-LGTHKPPANASDLTLGNEMPKYGYQGSWKLAWDPTANNCP 685  
DB 530 ENHD-LGKTQDPSFVQLSALGATATTTDVAAPTVAATP-THYGYQGTWGMVDDDTAST-P 586

QY 686 YTLKAT--WTKGYNPQPERVASLVNLSLWGSILDIRSAHSAIQASVDGRSVCRLGWVG 743  
DB 587 KTKTATLWNTGYLNPENRQGPLVPSNLMSGFSDIQAIGVTERALTLCSDRGFWAAG 646

QY 744 VSNFFYHRRDALQGGYRISGGYSL--GANSYFGSSMPGLAFTRVPGRSKYDYVCRSNHHA 802  
DB 647 VANFLDKDKGKRYKHKSGGYAIGGAAQTCSENILISPAFCQLFGSKDPLVAKNHTDT 706

QY 803 CIGSVYLS-----TQOALCGSYLPEDAFI--RASYGFQNGHMTSYFAESDV 849  
DB 707 YAGAFYIQTHTGCSFGTICLLDLKPLGWSHKLPLVLEQLAYSHSVNDLTKTYATPEVK 766

QY 850 RWNWNLACIGAGLPIVITPSKLY-LNELRPPVQAEFSYADHESFTTEEGDQARAFKSGH 908

```
Db 435 STLKQPVLTAGNLVKGVLDTKGTQTAGSSVIMDAGTTL-----KASTE 483
Qy 567 ITLSNLHLSLSLLANNAVTNPPNPAQDSHPAVIGSTTASGVITSGPIFFEDLDDTAY 626
Db 484 VTLTGLSIPVDSL-----GEGKKVITAASAASKVALSGPILLDDNQWY 529
Qy 627 DRYDWLGSQKINVLKLO-LGTKPPANAPSDLTLGNEMPKYQYQSWKLAWDPTANNGP 685
Db 530 ENHD-LGKTQDFSFVQLSALGATATTDPVAPTATP-THYGYQGTWGTWVDDTAST-P 586
Qy 686 YLKAT--WTKTYNPGPERSVSLVPSNLWGSILDIRSAHSAIQASVDGRSVCGLWVSG 743
Db 587 KTKATLAWTGYLDPNERQPLVPNSLWGSFSDIQAIOGVIERSAULTCSDRGFMAAG 646
Qy 744 VSNFFYHRRDALGQGYRISGYSI-LGANSYFGSSMFLAFTEVFGRSKDVVCRSNHA 802
Db 647 VANFLDKKKBKRYKHSKGYAIGGAATCSENLSIAPFQCLFGSKDPLVAKNHTDT 706
Qy 803 CIGSVLS-----TQALCGSYLFGDAFI--RASYGFGNQHMKTSYTFABESDV 849
Db 707 YAGAFYIQHTECSGFIQCLLDKLFSGWSHKPLVLEGQLAYSHVNDLTKYATPYEVKG 766
Qy 850 RWDNNCLAGEIGAGLPVITPSKLY-LNELRPPVQAEFSYADHESFTEEGDQARAFKSGH 908
Db 767 SWGNAPFNMQLQASHSY--PEYLHCFDYAPYIKNLITYIRQDSFSGTEGRSFDOSN 824
Qy 909 LLNLSPVGVKFDRCSTHFNKYSFMAAYICDAYRTISGTETTLTSHOETWTTDAFHAR 968
Db 825 LFNLSLPIGVAFKPEKFCNDSYDLTSLVVPDLIRNDPKCTALVISGASMETYANNLAR 884
Qy 969 HGTVRGWGYASLTNSINIEVGHGRYEYRDASRGYLSAGSRVP 1012
Db 885 QALQVRAGSHYAFSPMFVLOQFVPEVRSGSRIRYNVDLGGKQF 928

RESULT 8
D72077
polymorphic outer membrane protein g family - Chlamydomphila pneumoniae (strain CWL029)
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: D72077
R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A>Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:9920606; PMID:10192388
A:Accession: D72077
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <ARN>
A:Cross-references: UNIPROT:086164; UNIPARC:UPI000002PFEP; GB:AE001628; GB:AE001363; NID
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp 11
C:Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G

Query Match 20.7%; Score 1090; DB 2; Length 928;
Best Local Similarity 28.8%; Pred. No. 3e-58;
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;

Qy 1 MQTSFKHFLSNILAYSC--CSLGGGYAAIMVPGGIYDGETLVSF-----PYTVI 51
Db 1 MKTSPWLVSVSLAFSCHLOSAN-----BLLSPDSDFNIGNISGTFPTPKTSATVSLT 56
Qy 52 GDPSTTVFSAGELTKNLNLSIAALPL--SCFGNLGSFTVLGRGHSLTPTENIRTSNG 109
Db 57 GD--VPYFEPKGT-----PLSDSCFKQTTNLFNLGNHSLTGPIDAGTHA 102
Qy 110 AALGNSAADGLFTIEGFKELSFNSCNLSLAVLPAATYNGSQTPPTTSPNGTIYSKTD 169
Db 103 GAAASTANKNLTFSGFLLSFDs-----SPETTIVTQGTLLSAGG 144
Qy 170 LLLLNNEKFSYSNLVSGDGGDAIDAKSLTVQGISKLVQFQNTAQDGGACQVTSFSAM 229
```

```
Db 145 VNLNIRKLVVAGNFSTADGGAIKGAFLTGTSGDALFNNSSSTKGA----- 194
Qy 230 ANEAPIAFVANVAGVRGGIAAVODGGQGVSSSTSTEDPVVFSRNTA-VFEDGNVARVG 288
Db 195 -----IATTAGAR-----IANNITYVFLSNIAS 220
Qy 289 GGIYSYGNVAPLNNKTLPLNNVASPVYIAAKQPTSGOASNTSNNGYGGGAIFCKNGAQA 348
Db 221 GGAIDDEGTSILSNKFLYF-----EGNAAKTT-----GGAI-CNTRASG 259
Qy 349 G-----SNNGSVFSDGEGVVPFSSNVAAGKGGAIYAKLSVANCGPVPQIFLNTAND--- 400
Db 260 SPELIIENK-----TLIFASNVAFTSGGAHAKKALSGGTFEFLRNNVSSATP 310
Qy 401 -GGAIYLGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAISSMGSGKITTL 459
Db 311 KGGAISIDASGELSLSAETGNITFVRNLTIT--GSTD-----TPKRNAINIGSNKFTL 364
Qy 460 RAKAGHOILPNDPIEMANGNNQPAQSKLKLKINDEG-----YTGDIVFA----- 504
Db 365 RAAKNHTIFFYDPI-----TSEGTSDVLKINNGSAGALNPYQGTILFSGETILTABELK 418
Qy 505 ---NGSSTLYQNTIIEGRIVLREKAKLSVNSLSQTCGSLY-WEAGSTWDPVTPPPQP 560
Db 419 VADNLKSFQPVSLSGKLLQGVLTLESTFSQESQELGMDSGITLSTAGS----- 473
Qy 561 PAANQLITLSNLHLSLLANNAVTNPPNPAQDSHPAVIGSTTASGVITSGPIFFED 620
Db 474 -----ITITNLGINVDSLGLKQPV-----SILTAKGASKVKVIVSGKLNLD 513
Qy 621 LDDTAYDRYDWLGSQKINVLKLOGLTKPPANAPSDLTTLGNEMP-----KTYGQGSW 672
Db 514 IEGNIYESH-MFSDHQLFSLIKITV---DADVTNVDISSLIPVPAEDPNSEYGFQOQM 568
Qy 673 KLANDPNTANNGPYTLKATWTKGYNPGPERSVSLVPSNLWGSILDIRSAHSAIQASVDG 732
Db 569 NVNWTDTATN-TKEATATWTKGTFVSPSRKSAVLCNTLWGVFTDIRSLOQLVEIGATG 627
Qy 733 RSYCRGLWVSGVSNFFYHRRDALGQGYRISGYSILGANSYF-GSSMFLAFTVFGRSK 791
Db 628 MEHKQGVVSSMTNPLHKTGDNKRGFRHTSGGVVIGGSAHTPKDDLTPAFCHLPARDK 687
Qy 792 DYVVCNRNHHACIGSVYLSLTOQAL-----CGSYLFGDAFI-----RASYG 831
Db 688 DCFIAHNSRTYGTFLFPKSHSLTLPQNYLRGLRAKFSESAIEKFPREIPLALDVQVSFS 747
Qy 832 FGNQHKMSTYFAEESDVRWDDNCLAGEIGAGLPVITPSKLYLNLRLPFPVQAEPSYADH 891
Db 748 HSDNRMETHYTSLPESGSGWSNECIAGGIGLDLPVLSNPHPLPKTFIPQMKVEMVVSQ 807
Qy 892 ESFTEEGDQARAFKSGHLLMLSVPGVKFDRCSSTHFNKYSFMAAYICDAYRTISGTETT 951
Db 808 NSFFESSSDGSGFSGIGLLNLISIPVGAKFVQGDIGDSYTYDLSGFPFVSDVYRNNPQSTAT 867
Qy 952 LLSHOETWTTDAFHARHGVVVRGSMYASLTNSITVYGHGRYEYRDASRGYLSAGSRVR 1011
Db 868 LVMSPDWKIRGKGNLSRQAFLLRGSNNVYNSNCELFHYAMELGRSSRNTYNDVGTCLR 927
Qy 1012 P 1012
Db 928 P 928

RESULT 9
H86546
polymorphic outer membrane protein G family [imported] - Chlamydomphila pneumoniae (strain CWL029)
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86546
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A>Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
```

A:Reference number: A86491; MUID:20330349; PMID:10871362  
A:Accession: H8546  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-928 <STO>  
A:CROSS-references: UNIPROT:086164; UNIPARC:UPI000002FFBF; GB:BA000008; NID:g8978822; PI  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: pmp 11  
C:Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Query Match 20.7%; Score 1090; DB 2; Length 928;  
Best Local Similarity 28.8%; Pred. No. 3e-58;  
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;

QY 1 MTSFHKPFLSMILAYSC--CSLNGGYYAAEIMVPOGIYDGETLTVSF-----PYTVI 51  
DB 1 MKTSIPWLVSSVLAFLAFCHLQSLAN-----EELSPDDSFNGNIDSGTFTPKTSATYSLT 56

QY 52 GDPSTGVTSAGELTLKLDNSIAALPL--SCFNLGSGFTVLGRGHSITFENIRTSNG 109  
DB 57 GD---VFFYEPGKGT-----PLSDSCPKQTTDNLTLPLNGHSLTFFGIDAGTHA 102

QY 110 AALNSAAGDLFTIEGKELSPNCNLSLLAVLPAATNKGSOPTTTPSPNGTIYSKTD 169  
DB 103 GAAASTTANKNLTFSGFSLSPDS-----SPSTVTVTGQTLSSAGS 144

QY 170 LLLLNNEKFSFYENLVSGDGAIDAKSLTVQGISKLVCFOENTAQADGGACQVTSFSAM 229  
DB 145 VNLNIRKLVVAGNPFSTADGGAIKGASFLLTGTSGDALFNNSSSTKGA-----194

QY 230 ANEAPIAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTA-VEFDGNVARVG 288  
DB 195 -----IATTAGAR-----IANNTRYVRFSLNIASTS 220

QY 289 GGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASNTNNYGDGGAICKKGAQA 348  
DB 221 GGAIDDEGTSILSNKFLYP-----EGNAAKTT-----CGAI-CNTKASG 259

QY 349 G-----SNNSGVSFDGEGVFFSSNVAAGKGAIVAKKLSVANCQVPQFLRNAND--- 400  
DB 260 SPELIIISNNK-----TLIFASNVAETSGGAIHAKKALSGGGTPEFLRNVSATP 310

QY 401 -GGAIVLGESELSLSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAIISMGSGGKITL 459  
DB 311 KGAISIDASGELSLSAETGNITFVRNLTATT--GSTD-----TPKRNAINISNGKPTL 364

QY 460 RAKAGHOILPNDPLEMANGNQPQSSKLLKINDGEG-----YTGDIVPA----- 504  
DB 365 RAAKHHTIFFYDPI-----TSEGTSDVILKINNGSAGALNPVQGTILFSGEITLTADELK 418

QY 505 ---NGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTGGSLY-MEAGSTWDPVTPPQPQP 560  
DB 419 VADMLKSPFTQPSVLSGKLLQKGVLTSTSPQEAAGSLLGMDSGTTLSTTAGS----- 473

QY 561 PAANQLITLNLHLNLSLLANNAVTPNPPAQDHPAVIGSTTAGSVTISGPIPFED 620  
DB 474 -----ITITNLGINVDLSGLKQPV-----SLTAKGASNKIVSVGKMLNID 513

QY 621 LDDTAYDRYDLWLSGNQKINVLKLOCTKPPANAPSDLTNLNEMP-----KVGQGSW 672  
DB 514 IEGNIYESH-MPSHDQLFSLKITV-----DADVTNVDISLIPVPAEDPNSEYGFQGW 569

QY 673 KLAWDPNTANNPVTTLKATWTKGYNPQPERVASLVPNSLWMSILDIRSAHSATQASVDG 732  
DB 569 NVNNTTDTATN-TKEATATWTKTGTFVSPERKSALVCNLTNGVTDIRSLQOLVEIGATG 627

QY 733 RSYCRGLWVSGVSNFFYHNRDALQGVYRTISGGYSLGANSYP-GSSMFGLAFTVPRGSK 791  
DB 628 MEHQGFVSWMTNLFHKTGDENRKGFRHTSGGVYIGGSAHTPKDDLFTFAPCHLFARDK 687

QY 792 DYVVCRSNHHACISGVYLSLTOQAL-----CGSYLFGDAFI-----RASYG 831

DB 688 DCFIAHNNSRTYGGTLFFPKHSHTLOPQNYLRLGRAKFSESAIEKPPRIPLDALDVQVSFS 747  
QY 832 FGNQHMKTSYTFABESDVRWNNCLAGIGAGLPIVITPFSKLYLNLNLPFPVQAEFSYADH 891  
DB 748 HSDNEMETHYTSLPESEGSWNECIAGGIGLDLPVLSNPHPLPKTFIPQMKVEMVYVSQ 807  
QY 892 ESPTBEGDQARAPKSGHLLNLNLSVPVGVKDFRCSSSTPHNKYSFMAAYICDAYTIGTETT 951  
DB 808 NSFPSSSDGRGFSIGRLNLNLSIPVGAKFVQGDIGDSYTDLSGFFVSDVYRNPNQSTAT 867  
QY 952 LLSHQETWTTDAFLARHGVVVRGSMYASLTSNIEVGHGRVEYEDASRGYGLSAGSRVR 1011  
DB 868 LVMSPDNWKIRGNLSRQAFLLRGSNNVYNSNCLFGRYAMELGRSSRYNVVDVGTCLR 927  
QY 1012 F 1012  
DB 928 F 928

RESULT 10  
F81591  
polymorphic membrane protein G family CP0302 [imported] - Chlamydothila pneumoniae (str  
C:Species: Chlamydothila pneumoniae; Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 15-Jun-2001  
C:Accession: F81591  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: F81591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-949 <REA>  
A:CROSS-references: UNIPARC:UPI00001655FA; GB:AE002192; GB:AE002161; NID:g7189226; PIDN  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0302  
C:Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Query Match 20.7%; Score 1090; DB 2; Length 949;  
Best Local Similarity 28.8%; Pred. No. 3.1e-58;  
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;

QY 1 MTSFHKPFLSMILAYSC--CSLNGGYYAAEIMVPOGIYDGETLTVSF-----PYTVI 51  
DB 22 MKTSIPWLVSSVLAFLAFCHLQSLAN-----EELSPDDSFNGNIDSGTFTPKTSATYSLT 77

QY 52 GDPSTGVTSAGELTLKLDNSIAALPL--SCFNLGSGFTVLGRGHSITFENIRTSNG 109  
DB 78 GD---VFFYEPGKGT-----PLSDSCPKQTTDNLTLPLNGHSLTFFGIDAGTHA 123

QY 110 AALNSAAGDLFTIEGKELSPNCNLSLLAVLPAATNKGSOPTTTPSPNGTIYSKTD 169  
DB 124 GAAASTTANKNLTFSGFSLSPDS-----SPSTVTVTGQTLSSAGS 165

QY 170 LLLLNNEKFSFYENLVSGDGAIDAKSLTVQGISKLVCFOENTAQADGGACQVTSFSAM 229  
DB 166 VNLNIRKLVVAGNPFSTADGGAIKGASFLLTGTSGDALFNNSSSTKGA-----215

QY 230 ANEAPIAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTA-VEFDGNVARVG 288  
DB 216 -----IATTAGAR-----IANNTRYVRFSLNIASTS 241

QY 289 GGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASNTNNYGDGGAICKKGAQA 348  
DB 242 GGAIDDEGTSILSNKFLYP-----EGNAAKTT-----CGAI-CNTKASG 280

QY 349 G-----SNNSGVSFDGEGVFFSSNVAAGKGAIVAKKLSVANCQVPQFLRNAND--- 400  
DB 281 SPELIIISNNK-----TLIFASNVAETSGGAIHAKKALSGGGTPEFLRNVSATP 331

QY 401 -GGAIVLGESELSLSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAIISMGSGGKITL 459



A81591  
polymorphic membrane protein G family CP0307 [imported] - Chlamydomophila pneumoniae (strain  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: A81591  
C:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.P.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: A81591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-930 <REA>  
A:Cross-references: UNIPROT:Q92393; UNIPARC:UPI00001655PB; GB:AE002193; GB:AE002161; NID  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0307  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 20.6%; Score 1083; DB 2; Length 930;  
Best Local Similarity 28.4%; Pred. No. 8e-58;  
Matches 304; Conservative 160; Mismatches 406; Indels 202; Gaps 29;

QY 1 MTSFHKPFLSMILAYSCSLNGGGYAAEI-MVPGIYDGETLTVSPFVTIGDPSGTTV 59  
DB 1 MKIPLHKLISSTLV-TPILLSIATYGADASLSPDSDPAGGSGTTPKST-ADANGTNY 58  
QY 60 PSAGELTLKNDNSIAALPLSCFGNLLGSFTVLGRHSLTFENIRTSNGAALSAAADG 119  
DB 59 VLSGNNYINDAGKG-TALTGCCTTETGDLTGTGRGYSFNTVDAGSNAGAASATADK 117  
QY 120 LFTIEGFKELSPNCSNLSLAVLPAATNKGSTPTTT---STPSNGTIYSKTDLLLNNE 176  
DB 118 ALTTFTGFSNLSF-----IAAPGTTVASKGKSTLSAGALNLTNDGNTILFSONV----- 164  
QY 177 KFSFYNLSVGGGDAIDAKSLTVQGISKLCVQENTAQADGACQVVTFSAMANEAPTA 236  
DB 165 -----SNEANNNGGAIITKTLISGNTSITFTSNKAKLGGAIYSSAAASISGNTGQLV 219  
QY 237 FVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVEFDGNVARVGGIYSYGN 296  
DB 220 FMNN-KGETGGGALGFE-----ASSITQNSLFFSGNTATD----- 255  
QY 297 VAFLNNGKTLFLNNVASPVYIAAKOPTSQASNTSNNGYDGGAIKCKNGAQAAGSNNGSV 356  
DB 256 -----SHTIN-----QAFTQPLVVFVTAATAASDIYIDALLTSPVQTPPEHYGQHW 582  
QY 676 W-DPNTANNGPYTLKATWTKTGTPNGPGRVAVSLVPSNLGSLDILDIRSAHSAIQASVDGRS 734

A81591  
polymorphic membrane protein G family CP0307 [imported] - Chlamydomophila pneumoniae (stra  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: A81591  
C:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.P.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: A81591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-930 <REA>  
A:Cross-references: UNIPROT:Q92393; UNIPARC:UPI00001655PB; GB:AE002193; GB:AE002161; NID  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0307  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 20.6%; Score 1083; DB 2; Length 930;  
Best Local Similarity 28.4%; Pred. No. 8e-58;  
Matches 304; Conservative 160; Mismatches 406; Indels 202; Gaps 29;

QY 1 MTSFHKPFLSMILAYSCSLNGGGYAAEI-MVPGIYDGETLTVSPFVTIGDPSGTTV 59  
DB 1 MKIPLHKLISSTLV-TPILLSIATYGADASLSPDSDPAGGSGTTPKST-ADANGTNY 58  
QY 60 PSAGELTLKNDNSIAALPLSCFGNLLGSFTVLGRHSLTFENIRTSNGAALSAAADG 119  
DB 59 VLSGNNYINDAGKG-TALTGCCTTETGDLTGTGRGYSFNTVDAGSNAGAASATADK 117  
QY 120 LFTIEGFKELSPNCSNLSLAVLPAATNKGSTPTTT---STPSNGTIYSKTDLLLNNE 176  
DB 118 ALTTFTGFSNLSF-----IAAPGTTVASKGKSTLSAGALNLTNDGNTILFSONV----- 164  
QY 177 KFSFYNLSVGGGDAIDAKSLTVQGISKLCVQENTAQADGACQVVTFSAMANEAPTA 236  
DB 165 -----SNEANNNGGAIITKTLISGNTSITFTSNKAKLGGAIYSSAAASISGNTGQLV 219  
QY 237 FVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVEFDGNVARVGGIYSYGN 296  
DB 220 FMNN-KGETGGGALGFE-----ASSITQNSLFFSGNTATD----- 255  
QY 297 VAFLNNGKTLFLNNVASPVYIAAKOPTSQASNTSNNGYDGGAIKCKNGAQAAGSNNGSV 356  
DB 256 -----SHTIN-----QAFTQPLVVFVTAATAASDIYIDALLTSPVQTPPEHYGQHW 582  
QY 676 W-DPNTANNGPYTLKATWTKTGTPNGPGRVAVSLVPSNLGSLDILDIRSAHSAIQASVDGRS 734

583 WADTSTAKSG-----TMTVTTGYNPNPERRASVVPDLSWASFTDRTLQOIMTQANSIY 638  
QY 735 YCRGLWVSGVSNPFPHDRDALGCGQYRISGYSIGANSY-PGSSMFGLAFTVEFGRSKY 793  
DB 639 QORGLWASGTANFPHKDKSGTGNQAFRHKSIGYIVGGSAAEDSENIFFSVAFCQLFGKDKL 698  
QY 794 VVCSNHHACIGSVYLSSTQQAALCGSYL--FGDAF-----IRASYFGNQHMKTS 840  
DB 699 FIVENTSNNYLSLYLQHRAFLGGLPMPSPGSGITDMLKDIPILNAQLSYSYTKNDMDTR 758  
QY 841 YTFABESDVRWNNCLAGEIGAGLPIVITPSKLYLNELRPPVQAEFSVADHESFTEBDQ 900  
DB 759 YTSYPEAQSGWTNNSGALGGLSLALYLKPEAPFFQGYFPFLKFQAVYSROQNFKESGAE 818  
QY 901 ARAFKSGHLLNLSVPVGVKPRCSSTHKNYSFMAAICDAVRTISGTETLLSHQETWT 960  
DB 819 ARAFDDGDLVNCSPVGRILKSEDEKNNFEISLAYIGDVYRKNPRETSLSLWVGASWT 878  
QY 961 TDAFHARHGVVVRGSMYASLTNIYVGHGRYEDASRGYGLSAGSRVRF 1012  
DB 879 SLCKNLARQAFASAGSHLTLSPHVELSGEAAVELRGSAAHYNVDCGLRYSF 930

RESULT 13  
D72078  
polymorphic outer membrane protein g family - Chlamydomophila pneumoniae (strain CML029)  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: D72078  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.,  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: D72078  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-930 <ARN>  
A:Cross-references: UNIPROT:Q92393; UNIPARC:UPI000004708C; GB:AE001627; GB:AE001363; NID  
A:Experimental source: strain CML029  
C:Genetics:  
A:Gene: pmp\_8  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 20.5%; Score 1079; DB 2; Length 930;  
Best Local Similarity 28.3%; Pred. No. 1.4e-57;  
Matches 303; Conservative 160; Mismatches 407; Indels 202; Gaps 29;

QY 1 MTSFHKPFLSMILAYSCSLNGGGYAAEI-MVPGIYDGETLTVSPFVTIGDPSGTTV 59  
DB 1 MKIPLHKLISSTLV-TPILLSIATYGADASLSPDSDPAGGSGTTPKST-ADANGTNY 58  
QY 60 PSAGELTLKNDNSIAALPLSCFGNLLGSFTVLGRHSLTFENIRTSNGAALSAAADG 119  
DB 59 VLSGNNYINDAGKG-TALTGCCTTETGDLTGTGRGYSFNTVDAGSNAGAASATADK 117  
QY 120 LFTIEGFKELSPNCSNLSLAVLPAATNKGSTPTTT---STPSNGTIYSKTDLLLNNE 176  
DB 118 ALTTFTGFSNLSF-----IAAPGTTVASKGKSTLSAGALNLTNDGNTILFSONV----- 164  
QY 177 KFSFYNLSVGGGDAIDAKSLTVQGISKLCVQENTAQADGACQVVTFSAMANEAPTA 236  
DB 165 -----SNEANNNGGAIITKTLISGNTSITFTSNKAKLGGAIYSSAAASISGNTGQLV 219  
QY 237 FVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVEFDGNVARVGGIYSYGN 296  
DB 220 FMNN-KGETGGGALGFE-----ASSITQNSLFFSGNTATD----- 255  
QY 297 VAFLNNGKTLFLNNVASPVYIAAKOPTSQASNTSNNGYDGGAIKCKNGAQAAGSNNGSV 356  
DB 256 -----SHTIN-----QAFTQPLVVFVTAATAASDIYIDALLTSPVQTPPEHYGQHW 582  
QY 676 W-DPNTANNGPYTLKATWTKTGTPNGPGRVAVSLVPSNLGSLDILDIRSAHSAIQASVDGRS 734

```
Db 275 TISGNKSLTPAENS SVTGGCAICAGLDSRAAGPTLFNNRCNTAAAGKGAIAIADSGS 334
Qy 412 LLSLSADYGDIIIPDGNLKR TAKENAAADVGVTVSSQAISMGSGGKITTLIRAKAGHQILFND 471
Db 335 LLSLSANQGDITPLGN-----TLTSTAPTSTRNAIYLGSSAKITNLRAAQGQSIIYFD 387
Qy 472 PIEMANGN-----NQPAQSGKLKINDGEGYTGDIYPA-----NGSSTLY 511
Db 388 PIASNTTGASDLVLTINQPDNSPL-----DYSGTIVFSGEKLSDAKAADNFTSLK 440
Qy 512 QNVTTIEQGRIVLRERAKLSVNSLSOTGGSLVMEAGSTWDFVTPQPPQPPAANQLITLSN 571
Db 441 QPLASLSTLAKGNVLDVNGFTTEGSTLL-----MQGTGKLKADTEAISLTK 490
Qy 572 LHLSSLSLLANNAVTPNPTNPPAQDSHPAVIGSTTAG---SVTISGRPIFFEDLDDTAYDR 628
Db 491 LVVDLSALBGNKSV8-----IETAGANKTITLSPVFPQDSSGNPYE- 532
Qy 629 YDWLGSNOKINVLKQLQTKP-----PANAPSDLTLG-----NEMPKYGYQGSWKL 675
Db 533 -----SHTIN---QAFTQPLVWFVTAATAASDIYIDALLTSPVQTPBPHYGYQGHWEAT 582
Qy 676 W-DPNTANNPVTLKAATKTGYNPGRPRVSLVNSLWGSTLDIRSAHSAIQLASVDCRS 734
Db 583 WADTISKSG---TWMTVTTGYNPNERRASVWPSLWASFTDRTIQQIMTSQANSIY 638
Qy 735 YCRGLWVGVSNNFFYHRRDALQGGYRIYISGGYSLGANSY-FGSSMFGFLAFTEVFGRSXDY 793
Db 639 QORGLWASGTANFFHKDKSGTNQAPRHKS YGVIVGSAEDFSENI FSVAFQCLFGKDKDL 698
Qy 794 VVCRSNHACIGSVYLSLQQALCGSYL---FGDAP-----IRASYGFGNQHMTS 840
Db 699 FIVENTSHNYLASLYOHRALFGLGMPSPGSIITDMLKDIPILLNAQLSYSYTKNDMDTR 758
Qy 841 YTPAESDVRWDDNCLAGEIGAGLPVITPCKLYLNELRPVQAEFSYADHESFTEGDQ 900
Db 759 YTSYDEAGSMTNNGSALGELGSLALYLPKGAFFPGQFPFLKFOAVYSRQONPKESGAE 818
Qy 901 ARAFKSGHLLMLSVPGVKFORCSSTHPNKYSFMAAYICDAVRTISGTETTLISHQETWT 960
Db 819 ARAFDGDLVNCISIPVGIKLEKISDEKNNFEISLAYIGDVYRKNPRSRSTSLWVSGASWT 878
Qy 961 TDAFHARHVVVVRGSMYASLTSTNIEVYGHGRYEVDRASRGVGLSAGSRVRF 1012
Db 879 SLCKNLARQAFLASAGSHLTLSPHVELSGEAAVELRGSAAHYNVDCGLRYSF 930

RESULT 14
C72078
polymorphic outer membrane protein g family - Chlamydophila pneumoniae (strain CML029)
C:species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:accession: C72078
R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:reference number: A72000; MUID:99206606; PMID:10192388
A:accession: C72078
A:status: preliminary
A:molecule type: DNA
A:residues: 1-536 <ARN>
A:cross-references: UNIPROT:Q9Z988; UNIPARC:UPI00000470BC; GB:AE001627; GB:AE001363; NID
A:experimental source: strain CML029
C:genetics:
A:gene: pmp 7
C:superfamily: Chlamydophila pneumoniae polymorphic outer membrane protein G

Query Match 19.8%; Score 1041; DB 2; Length 936;
Best Local Similarity 29.0%; Pred. No. 2.9e-55;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

Qy 1 MQTSFHKFPLSNILAYSCSLNGGGYAAIMV--PQGIYDGETLTVSPFTVIGDPSGTT 58
```

```
Db 1 MKSVSMFLFSSIPFSSLSI-----VAEEVLDSSNNSYDGSNGTTFTVFPSTTDAAGTT 56
Qy 59 VFSAGELTNKLNDNSIAALPLSCFNLGSPVTLGRGHSITFFENIRT-STNGAALSNSAA 117
Db 57 YSLSDSVSPQNALGALPIPLASGCFLEAGGDLTTFQGNQHALKFAFINAGSSAGTVASTSA 116
Qy 118 DGLPTIRGFKELSFNSCNLSLAVLPAATNKGSTPTTTSTPSSNGTYISYKTDLLLLNNEK 177
Db 117 DKNLLFNDFSRLSIISCSPLIL-----SPT-----GQCALKSVGNLSLTGNSQ 159
Qy 178 FSPYSNLVSGDGGDAIDAKSLTVQGISKLCVQENTAQADGGACQVTVSFSAMANEAPATP 237
Db 160 IIFTQNFSSDGGGVINTKNFLLSCTSPASPSRNOA----- 195
Qy 238 VANTAGVRGGGIAAIVQDQGGVSSSTSTEDP-VVSFSRNTAVZPDGNVARVGGGIYSYGN 296
Db 196 ---FTGKGGVVYA-----TGTITIENSPOIIVSPQNLA---KGS---GGALYSTDN 238
Qy 297 VAFNNGKTLPLNNVSPVYIAAKOPTSGQASNTSNNYDGGGAIFCKNGAAGANNSSGV 356
Db 239 CSITDNFQVIFDGNSA---WEAAQ-----AOGGAICC-----TTTDTKV 274
Qy 357 SPDCGCVVPFSSNVAAGKGAIYAKLSVANGCPVQFLNINAND-----CGAIYLGESG 410
Db 275 TLTGKNLSPNTNALTGYGAIISGLKVSISAGGPTLFQSNISGSSAOGGGGGAINIASAG 334
Qy 411 ELSLSADYGDIIIPDGNLKR TAKENAAADVGVTVSSQAISMGSGGKITTLIRAKAGHQILFN 470
Db 335 ELALSATSGDITFNN-----QVTNGSTSTFNAINIIDTAKVTSIRATQSIYFY 385
Qy 471 DPIEMANGNQPAGSSKLLKINDGEG-----YTGDIVP-----ANGSSTLYQ 512
Db 386 DPIT---NPGTAASTDTLNLNLADANSEIYGAIVFSGEKLSPTEKATAAANTVTIRQ 441
Qy 513 NVTTIEQGRIVLRERAKLSVNSLSOTGGS-LYMEAGSTWDFVTPQPPQPPAANQLITLSN 571
Db 442 PAVLARGDLVLDGVVTVFTKDLTQSPGSRILMDGGTTLS-----AKEANLSLNG 490
Qy 572 LHLSSLSLLANNAVTPNPTNPPAQDSHPAVIGSTTAGSVTISGTETTLISHQETWT 631
Db 491 LAVNLSSLDGTNKA-----ALKTERADKNISLSGTIALIDTEGSPVENH- 535
Qy 632 LGSNOKINVLKQLQGTKPPAN-----APSDLTIGNEMPKYGYQGSWKLAWDPNTANNP 685
Db 536 LKSASTYPLLEL---TTAGANGTTITLGAALSTLTQEPETHYGYQGNQWLSWANATSSK-- 590
Qy 686 YTLKATWTKGYNPGRPRVSLVNSLWGSTLDIRSAHSAIQLASVDCRSYCRGLWVGVS 745
Db 591 -IGSINWTRGTGYPSPRKNPLNLSLWGNPDIRSINQLIETKSGSEPPFERELWLSGIA 649
Qy 746 NFFYHRRDALQGGYRIYISGGYSLGANSYF-GSSMFGFLAFTEVFGRSKDYVVCRRNHACI 804
Db 650 NFFYRDSMPTREHGFHISGGYALGITATPAEDQLTFAPCOLFARDNRNHIITGNHGDYTG 709
Qy 805 GSVYLSLQQAL--CGSYLFGDAP-----IRASYGFGNQHMTSYTFA 844
Db 710 ASLYFHFHTEGLFDIANFLWKGATRAPWVLSISQIIPLSFPAKFSYLTDNHNKTYTT-- 767
Qy 845 EESDVR--WDNNCLAGBIGAGLPVITPCKLYLNELRPVQAEFSYADHESFTEGDQAR 902
Db 768 DNSIIKGSWRNDAPCADLGASLPVIVSVFY-LKEVEFPVKVQIYIAHQDQFYERHAEGR 826
Qy 903 AFKSGHLLNLSVPVGVKFORCSSTHPNKYSFMAAYICDAVRTISGTETTLISHQETWT 962
Db 827 AFNKSSELINVEIPGVTFPERDSKSEKGYDILTLMYILDAYERNPKCQTSLSIADANMAY 886
Qy 963 AFHLARHGVVVRGSMYASLTSTNIEVYGHGRYEVDRASRGVGLSAGSRVRF 1012
Db 887 GTNLARQGSVRAAHNFQVNPHEIFQGFAPFVSSSRNTNINLGSKFCF 936
```

RESULT 15



B81591  
polymorphic membrane protein G family CP0308 [imported] - Chlamydomophila pneumoniae (stra  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: B81591  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: B81500; PMID:20150255; PMID:10684935  
A:Accession: B81591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-936 <REA>  
A:Cross-references: UNIPROT:Q92898; UNIPARC:UPI0000131CED; GB:AE002161; NID  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0308  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 19.7%; Score 1040; DB 2; Length 936;  
Best Local Similarity 29.0%; Pred. No. 3.3e-55;  
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

QY 1 MQTSFKPFLSMILAYSCSLNGGGYAAEIMV--PQGIYDGETLTVSPFTVIGDPSGTT 58  
DB 1 MKSSVSWLPFSSIPLFSSLSI-----VAAEVLDSNNNSVDGSGNGTTFVFTTDDAAGTT 56

QY 59 VFSAGBELTKNLNDNSTAALPLSCFGLNLSFTVLGRHSLTFENIRT-STNGGAALNSAA 117  
DB 57 YSLSDVSQFQACALGIPLAGSGCFLEAGDGLTFQGNQHALKFAFINAGSAGTVASTSA 116

QY 118 DGLFTIEGFKELSFNSCNLSLAVLPAAATNKGSGTPTTTPSNGTIYSKTDLLLNNEK 177  
DB 117 DKNLFPNDPRLGIISCPSSL-----SPT-----GQCALKSQVNLSTGNSQ 159

QY 178 FSYNLVSDGGAIDAKSLTVQISKLCVFOENTQAADGACQVVTSPSAMANEPIAF 237  
DB 160 IITQNFSSDNGGVINTKNFLSGTQSPASFSRQA----- 195

QY 238 VANVAVRGGGIAAODGQGVSSSTSTEDP-VVFSRNTAVFDGNVARVGGGIYSYGN 296  
DB 196 ---FTGQGGVVYA-----TGTITIENSPGIVSFSQNLA---KGS---GGALYSTDN 238

QY 297 VAFNLNGKTLFLNNVASPVVIAAKQPTSGQASNTNNYDGGAIKCKNGQAQAGSNNSGV 356  
DB 239 CSITDNFQVIFDGNSA---WEEAQ-----AQQGAICC-----TTTDKTV 274

QY 357 SFDGEGVVFFSSNVAAGKGAIVAKKLSVANCGVPQFLRNIAND-----GGAIYLGESG 410

591 -IGSNWTRTGYIPSPERKSNLPLNSLWGNFDIRSINQLIETKSSGPPFRELWLSGIA 649  
746 NFFVHDDRALCOQVRYISGGYSLGANSYF-GSSMFGLAFTVFGRSKDYVWVCRNHHACI 804  
650 NFFYRDSMPTRHGFHISGGYALGATATPAEDQLTFAPCQLFADRHHITCKNHDGYG 709  
805 GSVLSTOQAL--CGSYLFGDAF-----IRASYGFGNQMKTSYTPA 844  
710 ASLYPHHTEGLFDIANFLWKGATAPWVLSBISQIIPLSFQAKFSYLTNDHMKTYT-- 767  
845 RESDVR--WDNNCLAGEIGAGLPVITPSKLYLNLBPVQAFSPYADHESPTBEGDQAR 902  
768 DNSIIKGSWRNDAFACADLGASLPFVISPYL-LKEVEFPVKVQYIYAHQODFYERYAER 826  
903 AFKSGHLANLSPVGVKDRCSSTHPNKYSFMAAYICDAYRTISCTETTLTSHQETWTTD 962  
827 AFNKSELINVELPIGVTFERSKSEKGYDILTMILDAYRENPKQTSLSIADANWMA 886  
963 AFHLARHGVVVRGSMYASLTSNIEVYHGRVEYRDASRGYGLSAGSRVRF 1012  
887 CTNLRQGSVRAANHFOVNPHEIFGQFAPEVRSSRNNTNLGSKFCF 936

RESULT 16  
C86546  
polymorphic outer membrane protein G family [imported] - Chlamydomophila pneumoniae (stra  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: C86546  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I-  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A:Reference number: A86491; PMID:20330349; PMID:10871362  
A:Accession: C86546  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-936 <STO>  
A:Cross-references: UNIPROT:Q92898; UNIPARC:UPI0000131CED; GB:BA000008; NID:98978817; P  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: pmp 7  
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 19.7%; Score 1040; DB 2; Length 936;  
Best Local Similarity 29.0%; Pred. No. 3.3e-55;  
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

QY 1 MQTSFKPFLSMILAYSCSLNGGGYAAEIMV--PQGIYDGETLTVSPFTVIGDPSGTT 58  
DB 1 MKSSVSWLPFSSIPLFSSLSI-----VAAEVLDSNNNSVDGSGNGTTFVFTTDDAAGTT 56

QY 59 VFSAGBELTKNLNDNSTAALPLSCFGLNLSFTVLGRHSLTFENIRT-STNGGAALNSAA 117  
DB 57 YSLSDVSQFQACALGIPLAGSGCFLEAGDGLTFQGNQHALKFAFINAGSAGTVASTSA 116

QY 118 DGLFTIEGFKELSFNSCNLSLAVLPAAATNKGSGTPTTTPSNGTIYSKTDLLLNNEK 177  
DB 117 DKNLFPNDPRLGIISCPSSL-----SPT-----GQCALKSQVNLSTGNSQ 159

QY 178 FSYNLVSDGGAIDAKSLTVQISKLCVFOENTQAADGACQVVTSPSAMANEPIAF 237  
DB 160 IITQNFSSDNGGVINTKNFLSGTQSPASFSRQA----- 195

QY 238 VANVAVRGGGIAAODGQGVSSSTSTEDP-VVFSRNTAVFDGNVARVGGGIYSYGN 296  
DB 196 ---FTGQGGVVYA-----TGTITIENSPGIVSFSQNLA---KGS---GGALYSTDN 238

QY 297 VAFNLNGKTLFLNNVASPVVIAAKQPTSGQASNTNNYDGGAIKCKNGQAQAGSNNSGV 356  
DB 239 CSITDNFQVIFDGNSA---WEEAQ-----AQQGAICC-----TTTDKTV 274

QY 357 SFDGEGVVFFSSNVAAGKGAIVAKKLSVANCGVPQFLRNIAND-----GGAIYLGESG 410





Db 265 LALTGNTQVLFOENKTTGSAQAANNPEGCGGAIACCYLATATDKTGLAISQNEMSPTSNT 324  
Qy 185 VSGDGAIDAKSLTVQGISKLCVFOENTAQAD-GGACQVVT-SFSAMANEAPAFVANVA 242  
Db 325 TTANGGAIYATKCTLDGNTTL-TFDQNTATAGCGAIYTETEDFSLKSGTGTVTFSTNTA 383  
Qy 243 GVRGGI-----AAVDGQGV-----SSSTST----- 265  
Db 384 KT-GGALYSKGNSSLTGTNLLFSGNKATGPSNSANQEGCGGAILSFLESASVSTKGL 442  
Qy 266 ---EDPVVFSRNTA-----VEFDGNVAR-VGGGIY----- 292  
Db 443 WIEDENVSLSGNTATVSGGAIYATKCALHGNNTLTFDGNATAGGAIYTETEDFTLTG 502  
Qy 293 SYGNVAFNN-----GKTLFLNNVASPVYIAAKOPTSQOASNTSNNGDGAIF 341  
Db 503 STGTVTFTNTAKTAGALHTKNTSFTKKA---LVPSGNSATATATTTTQEGCGGAIL 559  
Qy 342 C-----KNGAQAGNSNGSVS-----FDGEGVVPFSSNVAAGKG 376  
Db 560 CNISESDIATKSLTLTENESLSFINNTAKRGGGIYAPKCVISGESINFQDNTAETSGG 619  
Qy 377 AIYAKKLSVANCGPVQFLRNITANDGGAIYLGSELSADYGDIIIPDGNLKRKTAKENAA 436  
Db 620 AIYSKNLSITANGPVSFNNSGGKGAIIYADSGELSLEADGDITPSGN---RATEGTS 676  
Qy 437 DVNGVTSSQAI SMGSGGKIITLRAKAGHOILFNDPIEM---ANG----- 478  
Db 677 TPN-----SIHLGAGAKITKLAAPGHTIYFYDPITMEAPASGGTIEELVINPVVKAI 729  
Qy 479 -----NNQPAQSKLKLKINDGEGYTDIVP-----ANGSSTLYQNVTIEQG 519  
Db 730 VPPQPKNGPIASVPVVPVAPANPNTGTIVFSSGKLPSQDASIPANTTTILNQKINLAGG 789  
Qy 520 RIVLRKAKLSVNSLSQGGG-LYMEAGSTWDFVTPQPPQPPAANQILITSLNHLSS 578  
Db 790 NVVLKEGATLVQSFQPDSTVPMDAGTTLETITTN-----NTDGSIDLKNSVLDA 843  
Qy 579 LLANNAVTPNPPAQDSHPAVITSGPTAGSVTISGPIFPEDLDDTAYDRYDWTGNSQKI 638  
Db 844 LDGKRMIT-----IAVNSTSGGLKISGDLKPHNNGSVFYDN---PGLKANL 886  
Qy 639 NVLKLQL-GTKPPAN-----APSDLTLGNEMPKYQGSWKLANDPNTANNGPYTLKAT 691  
Db 887 NLPFLDLSSSTGTVNLDNFNIPSSMA-----APDYGQGSWTLV---PKVGAGGKVTLVAE 940  
Qy 692 WTKGTGYNPGRVASLVNSLWGSTLDIRSAHSAIQAASVDCRSYCRGLWVSGVSNFFVHD 751  
Db 941 WQALGYTPKPELRATLVNSLWVNYVTHSQEIIATAMSDAPSHPGIWIQIGIAGAFQD 1000  
Qy 752 RDALGQGVYISGGYSLGANSYFGSS-MFGLAFTVFGRSKDYVYCRSNHHACIGSVYLS 810  
Db 1001 KOKENAGPRLISRGYIVGSGMTTPOEYTFVAFBQLFGSKDYVYSDIKSQVYAGS--- 1056  
Qy 811 TQQAIC--GSYLP-----GDAPI-----RASYGFGNQMKTSYTAE 845  
Db 1057 ----LCAQSSYVIFLHSLRRLHVLSKVLPELPGTPLVLHGQVSYGRHHNMITKLANNT 1112  
Qy 846 ESDVRWNNCLAGIETAGLPITVITPSKLYLNLAPFPVQAEFSYADHESFTBEGDQAAFK 905  
Db 1113 QKSDWDHSHFAVEVGSLPVDL--NRYLTSYSPYVKLVQVSVNQKQFQVAADPRIFD 1170  
Qy 906 SGHLNLNVPGVGFRCSSSTHPNKYSFMAIYICDARTISGTETTLTLLSHOETWTTDAFH 965  
Db 1171 ASHLVNSVSI PMGLTFKHESAKPPSALLLLTGYADVADYDRHPCLTS-LTNGTSSWTFATN 1229  
Qy 966 LARGVVRGMYASLTNSNIEVYGHGRYVEDASRGYGLSAGSRVF 1012  
Db 1230 LSROAFPAASGHLKLHLGLDCFASGCELSRSSRSRYNANCGRYSF 1276

RESULT 20  
C81591

polymorphic membrane protein G family CP0309 [imported] - Chlamydomophila pneumoniae (stra  
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C;Accession: C81591  
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A;Reference number: A81500; MUID:20150255; PMID:10684935  
A;Accession: C81591  
A;Status: preliminary  
A;Molecule type: DNA  
A;Releaves: 1-1276 <REA>  
A;Cross-references: UNIPROT:Q92899; UNIPARC:UPI0000131CEC; GB:AE002161; NID  
A;Experimental source: strain AR39, HL cells  
C;Genetics:  
A;Gene: CP0309

Query Match 19.2%; Score 1012; DB 2; Length 1276;  
Best Local Similarity 25.9%; Pred. No. 2.6e-53;  
Matches 339; Conservative 153; Mismatches 417; Indels 398; Gaps 40;  
Qy 38 DGETLTVSPVTVIGDPSTTVFSGRLTLKNDNSTAALPLSCFGLNLSFTVLGRGHS 97  
Db 36 NGSGSAAPFAKETSDASGTYTLTSDVSTNV-SAITPADKSCFTTGGALSFPVGADHS 94  
Qy 98 LTPENIRTSNGALSNAADGLFTTREGPKELSPNCNSLL----- 138  
Db 95 LVLQTLIALTHDGAANNNT-----ALSPFGSFLILDSAPATGTSGGKGAICVT 144  
Qy 139 -----AVLPA 143  
Db 145 NTEGATFTDNASVTLQNTSEKGAASAYSIDLAKTTTAAALLDQNTSTKNGALCST 204  
Qy 144 ATT-----KNGSQT-PTTSTPSNGTIYS----- 166  
Db 205 ANTVQNGSVTVTFSSNTATDKGGIYSKEKSDLDANTGVTVFKNSKTATKGNWSSDDN 264  
Qy 167 -----KTDLLLNNEKPSFYSNL 184  
Db 265 LALTGNTQVLFOENKTTGSAQAANNPEGCGGAIACCYLATATDKTGLAISQNEMSPTSNT 324  
Qy 185 VSGDGAIDAKSLTVQGISKLCVFOENTAQAD-GGACQVVT-SFSAMANEAPAFVANVA 242  
Db 325 TTANGGAIYATKCTLDGNTTL-TFDQNTATAGCGAIYTETEDFSLKSGTGTVTFSTNTA 383  
Qy 243 GVRGGI-----AAVDGQGV-----SSSTST----- 265  
Db 384 KT-GGALYSKGNSSLTGTNLLFSGNKATGPSNSANQEGCGGAILSFLESASVSTKGL 442  
Qy 266 ---EDPVVFSRNTA-----VEFDGNVAR-VGGGIY----- 292  
Db 443 WIEDENVSLSGNTATVSGGAIYATKCALHGNNTLTFDGNATAGGAIYTETEDFTLTG 502  
Qy 293 SYGNVAFNN-----GKTLFLNNVASPVYIAAKOPTSQOASNTSNNGDGAIF 341  
Db 503 STGTVTFTNTAKTAGALHTKNTSFTKKA---LVPSGNSATATATTTTQEGCGGAIL 559  
Qy 342 C-----KNGAQAGNSNGSVS-----FDGEGVVPFSSNVAAGKG 376  
Db 560 CNISESDIATKSLTLTENESLSFINNTAKRGGGIYAPKCVISGESINFQDNTAETSGG 619  
Qy 377 AIYAKKLSVANCGPVQFLRNITANDGGAIYLGSELSADYGDIIIPDGNLKRKTAKENAA 436  
Db 620 AIYSKNLSITANGPVSFNNSGGKGAIIYADSGELSLEADGDITPSGN---RATEGTS 676  
Qy 437 DVNGVTSSQAI SMGSGGKIITLRAKAGHOILFNDPIEM---ANG----- 478  
Db 677 TPN-----SIHLGAGAKITKLAAPGHTIYFYDPITMEAPASGGTIEELVINPVVKAI 729  
Qy 479 -----NNQPAQSKLKLKINDGEGYTDIVP-----ANGSSTLYQNVTIEQG 519  
Db 730 VPPQPKNGPIASVPVVPVAPANPNTGTIVFSSGKLPSQDASIPANTTTILNQKINLAGG 789

[illegible]

Qy	209	QENTAOADGGACQCVTTSFPMANEAPIAFVANVAGVYRGGLIAAVQDQCGVSSSTSTEDP	268
Db	584	SGNTATVSGGA---IYATKCALHGNTTLLTDFGNTAETAGGAI-----YTETEDP	629
Qy	269	VWFSRNTAVFEDGNVARVGGGIYSYGNVAFLMNGKTLFLNNVASPVYIAAKQPTSGQAS	328
Db	630	TLTGSTCT-VTFSTNTAKTAGALHTKGNTSFTKQKALVFGNSA-----TATAT	677
Qy	329	NTSNVYDGGAIFC-----KNGAQAGSNNSGVS-----PDGEGV	363
Db	678	TTTDQEGCGGAIICNTSESIDATKSTLTENESLSFINNTAKSGGGIYAPKCVISGES	737
Qy	364	VFSSNVAAKGGAIIYAKKLIVANCGPVPQFLRIANDGGAIYLGSEGLSLSADYDIIIF	423
Db	738	INFDCNTAETSGGAIYKNSLITANGPVSFTNNSGKGGAIIYADSGELSLEAIDGDITF	797
Qy	424	DGNLRTAKENAADVNGVTVSSQAISNGSGKKITTLRAKAGHOILFNDDPIEM---ANG--	478
Db	798	SGN---RATEGTSTPN-----SIHLGAGAKITKLAAAPGHTIYFYDPTIMEAPASGGT	847
Qy	479	-----NNQPAQSSKLLKINDGEGYTGDIYF-----ANG	506
Db	848	IEELVINPVVKAIVPPQPKNGPIASVPVVPVAPAMPNTGTIIVFSSGKLPSQDASIPANT	907
Qy	507	SSTLYQNVTTBOGRIVLREKAKLSVNSLSQTGGS-LYMEAGSXTWDFVTPOPPQPPAANQ	565
Db	908	TTTILNQKINLAGNVVLKEGATLQVYSFTQOPDSTVPMDDAGTTLTETTTN-----NTDG	961
Qy	566	LITLSNLHLSSLLANNAVTPNPTPPAQDSHPAVIGSTAGSVTISGPIFFEDLDDTA	625
Db	962	SIDLKSLSVNLDALDGRMIT-----IATNSTGGLTKISGLDKLPHNNNGSF	1007
Qy	626	YDRYDMLGSKNKNVLKQL-GTKPPAN-----APSDLTLGNEMPKYQGSGWKLAWDP	678
Db	1008	YDN---PELKANLMLPFLDLSSTSGTIVNLDLFPNIPSSMA-----APDYGYQGSWTLV--P	1058
Qy	679	NTANNGPYTLKATTKTKGYNPGPERRVASLVPSNLWGSILDIRSAHSAIQASVDGRSVCRG	738
Db	1059	KVGAGGKVTLVAEQALGYTPKPELRATLVPSNLWNAVYNIHSTIQOBIATAMSDAPSHPG	1118
Qy	739	LWVSGVSNFFHDDRALCOGVRYISGCVSGANSYFGSS-MFLGAFTEVFGRSKDYVCCR	797
Db	1119	IWIGIGNAFHODKOKENAGPFLISRGYIVGGSMTTPOEYTFVAFVSQLFGKSKDYVVSD	1178
Qy	798	SNHACIGSVYLSLTOQALC--GSYLF-----GDFAI-----RASVGF	832
Db	1179	IKSQVYAGS-----LCAQSSYVILPHSLRRLHVLKVLPELPGETPLVLHGQVSYGR	1230
Qy	833	GNQHKMSTYTPABSDVRWDNNCLAGEIAGLPIVITPSKLYLNELPFPVQAEFSYADHE	892
Db	1231	NHNMTTKLANNTQKSDWDSDSHSAVEVGGSLPVDL--NYRYLTSYFPYVKLQVIVSNQK	1288
Qy	893	SFTEGDQARAFKSHLLNLSPVGVKFDRCSSVTHPNKYSFMAAYICDARTISGTETTL	952
Db	1289	GFQEVAAOPRIFDASHLVNVSIPMGLTYPKHESAKPPSALLTLTLGYADVADYRDHPCLTS-	1347
Qy	953	LSHQETWTTTDAFLARHGVVVRWGSYASLTSNIENVYGHGREYERDASRGYGLSAGSRVRF	1012
Db	1348	TYGCTSWCTETNTI-SFOAPPAPASGHIKLIHGLDCPAGSGCBLSSRSRNYNANGCTYSF	1407

	Query Match	18.7%;	Score	985.5;	DB 2;	Length	1407;			
	Best Local Similarity	28.0%;	Pred. NO.	1.2e-51;						
	Matches	302;	Conservative	148;	Mismatches	404;	Indels	226;	Gaps	34
Qy	51	IGDSPGTTVFSAGELTLKLNLSI---	AALPLSCFGNLLGSFTVLGRHSITFFENIRATST	107						
		: : : : : : : : : : : :	:	:	:	:	:	:	:	
		: : : : : : : : : : : :	:	:	:	:	:	:	:	
		: : : : : : : : : : : :	:	:	:	:	:	:	:	
		: : : : : : : : : : : :	:	:	:	:	:	:	:	
		: : : : : : : : : : : :	:	:	:	:	:	:	:	
		: : : : : : : : : : : :	:	:	:	:	:	:	:	



Db 600 LMGNAVDVRSMLMQVQTHASSLQTDRLGLWDIGNFFHVSAEEDNIRYHNSGGYVLSVN 659

QY 772 SYFGSSMP-GLAPTEVFGSKDYVVCVSRNHACIGSVYLSLQALCGSYLF-----GDA 824

Db 660 NEITPKHTYMAFSQLFSDKDYAVSNNEIR-----MYL-----GSYLYQYTTSLGNI 707

QY 825 FIRAS-----YGFQNHMKTSYTFABESDVRWNNCLIA 857

Db 708 FRYASRNPVNVGILSRRLQNLPLMFHFLCAYGATNDKTDYANFPMVKNSWRNCCWA 767

QY 858 GBIGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTTEGQDQARAFKSGHLLNLNSVPVG 917

Db 768 IECGSMPLLVFENGRLFQCAIPEFMKQLVYAYQGDFTKTTADGRFSGSLTSISVPLG 827

QY 918 VKFDRCSSTHPNKYSFMAIYICDARTISGTETLLSHQETWTTDAFLARHGUVVVRGSM 977

Db 828 IRPEKLALSQDVLVDFSPSYIPIFRKOPSCAAALVISGDSMLVPAAHVSRHAFVGSGRG 887

QY 978 YASLTSNIENVYGHGREYRDASRGYLSAGSRVRF 1012

Db 888 RYHFNDDYTELLCGSIECPHARNYINCSKFRP 922

RESULT 24

E86491

polymorphic outer membrane protein G family [imported] - Chlamydia pneumoniae (strain E86491)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E86491

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Inoue, Y.

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349; PMID:10871362

A:Accession: E86491

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-922 <STO>

A:Cross-references: UNIPROT:Q9Z9G5; UNIPARC:UPI000004708A; GB:BA0000008; NID:g8978378; PMID:10871362

A:Experimental source: strain J138

C:Genetics:

A:Gene: pmp.1

C:Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Query Match 18.4%; Score 967; DB 2; Length 922;

Best Local Similarity 28.0%; Pred. No. 8.9e-51;

Matches 295; Conservative 144; Mismatches 380; Indels 236; Gaps 32;

QY 31 MYPQGIYDGTTLTVSPFYTVIGDPSTVPSAGELTKNLNDNSIAALPLSCFQNLGSGFT 90

Db 31 LTPEDSFHGDSONAERSYV---QAGDVYSLTGDVSIINVDNS--ALNKACFNVTSGSVT 85

QY 91 VLGRGHSLTFENITST--NGAALSNAADGLFTIEGFKELSFNSNCNLSLLAVLPAATTNK 148

Db 86 FAGNHGLYFNINSSGTTKEGAVLCCQDPQATARFSGFTLSF----- 128

QY 149 GSQPTPTTSPNGTIYKTDLLLNKESFVSNLVSGGGDAKSLTVQIGSKLCPV 208

Db 129 -IQSP--GDIKEQCLYSKRNALLNYYVRFQNSQTKKGAIISGANVTIVNGSVSP 185

QY 209 QENTAQADGCGCQVTSFSAANEAPAFVANVAGVRGGGIAAVQDQCGQVSSSTSTEDP 268

Db 186 YQNAATFG-----ALHSGGLQIAVQAEIR----- 212

QY 269 VVPSRNTAVFEDGNVARGGIIYSGNVAFNLNNGKTLFLNNVASPVYIAAKQPTSGQAS 328

Db 213 ---PAQNTAK--NGS-----GGALYSDGDIDIDQNAVYLFRENEALTAI----- 252

QY 329 NTSNNYGGGAIFCKNGAQAGSNNS-----VSPDGEVGVFPSSNVAAKGGAIIYAKK 382

Db 253 -----GKGGAIVC-----LPTSGSSTPVPVITFSDNKQLVPERHNSIMGGAIIYARK 299

QY 383 LSVANCGPVQFLRNIA-----NDGAIYLGESGELSLADYGDIIIPGNLKRKTAKENAD 437

Db 300 LSISSGGPTLFINNISVANSQNLGAIAIDTGGGIISSAEKGTITPQGN-----RTSLPP 354

QY 438 VNGVTSSQALSMGSGKITTTRAKAGHQIILFNDPIEMANGNNOQAOSKLLKIN---DG 494

Db 355 LNHILLQNA-----KFLKLAQARNGYSIEFYDPI-----TSEADGSTQLNNGDPKN 401

QY 495 EGYTGDIYVA-----NGSSTLYQNVITBOGRIVLRKAKLSVNSLSQTTGGS-LYM 543

Db 402 KEYTGTILFSGEKSILANDPRDFKSTIPQNVNLSAGLYIKEGAERYVSKFTQSPGSHLV 461

QY 544 EAGSTWDFVTPQPPQAPAAQNLITLSNLHLSSLANNANVTNPTTPPAQDSHPAVIG 603

Db 462 DLGT-----KLITASKEDITAITGLAIDIDL-----SSSTTAATVTK 496

QY 604 STTAG---SVTISGPIF-----FEDLDDTAYDYDWLGSNQKINVLKQLGTKPPANAP 654

Db 497 ANTANKQISVTDSTIELISPTGNAYEDLR-----MENSQTFFLLSLE-----PGAGGS 543

QY 655 SDLTILGNEM---PKYGYQGSWKLAWDPNTANNPPTLKATWTKTGYNPGPERSVASLVPNS 711

Db 544 VTVTAGDPLVPSPHYFGQGNWKLAWTGTGNKVGEP---FMDKINYPKPEKEGNLVPNI 599

QY 712 LMGSTILDIRSAHSATQASVDGSRGCRGLWVSGVSNFFYHRRDALGGVRYIISGGYSLGAN 771

Db 600 LMGNAVDVRSMLMQVQTHASSLQTDRLGLWDIGNFFHVSAEEDNIRYHNSGGYVLSVN 659

QY 772 SYFGSSMP-GLAPTEVFGSKDYVVCVSRNHACIGSVYLSLQALCGSYLF-----GDA 824

Db 660 NEITPKHTYMAFSQLFSDKDYAVSNNEIR-----MYL-----GSYLYQYTTSLGNI 707

QY 825 FIRAS-----YGFQNHMKTSYTFABESDVRWNNCLIA 857

Db 708 FRYASRNPVNVGILSRRLQNLPLMFHFLCAYGATNDKTDYANFPMVKNSWRNCCWA 767

QY 858 GBIGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTTEGQDQARAFKSGHLLNLNSVPVG 917

Db 768 IECGSMPLLVFENGRLFQCAIPEFMKQLVYAYQGDFTKTTADGRFSGSLTSISVPLG 827

QY 918 VKFDRCSSTHPNKYSFMAIYICDARTISGTETLLSHQETWTTDAFLARHGUVVVRGSM 977

Db 828 IRPEKLALSQDVLVDFSPSYIPIFRKOPSCAAALVISGDSMLVPAAHVSRHAFVGSGRG 887

QY 978 YASLTSNIENVYGHGREYRDASRGYLSAGSRVRF 1012

Db 888 RYHFNDDYTELLCGSIECPHARNYINCSKFRP 922

RESULT 25

F81539

polymorphic membrane protein G family CP0770 [imported] - Chlamydia pneumoniae (strain F81539)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 15-Jun-2001

C:Accession: F81539

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, M.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.L.; et al.

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: F81539

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-922 <REA>

A:Cross-references: UNIPARC:UPI000016560D; GB:AE002237; GB:AE002161; NID:g7189684; PIDN:10684935

A:Experimental source: strain AR39, HL cells

C:Genetics:

C:Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Query Match 18.3%; Score 966; DB 2; Length 922;

Best Local Similarity 28.0%; Pred. No. 1e-50;

Matches 295; Conservative 143; Mismatches 381; Indels 236; Gaps 32;

QY 31 MYPQGIYDGTTLTVSPFYTVIGDPSTVPSAGELTKNLNDNSIAALPLSCFQNLGSGFT 90



```
Db 31 LTPEDSFHGDQNAESYV---QAGDYVSLTGDVSIQNVDS--ALNKAQFNTVSGSVT 85
Qy 91 VLGRGHSUTFTNIRST--NGAALSNSAADGLFTTIEGPKELSPFNCNLSLLAVLPAATNKK 148
Db 86 FAGNHGGLYFNNISSGTTKEGAVLCCQDQATAREPFGFTLSF-----128
Qy 149 GSQTPPTTSTPBGNYSTKTDLLILANNEKFPSPYSLVSGDGAIDAKSLTVQGISKLCVP 208
Db 129 -IQSP--GDIKEQGCLYSKNALMLLNYYVVRPEQNSKTGKGAIKGANVTIVGNYDSVSF 185
Qy 209 QONTAAQADGACQVVTFSAMANEAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDP 268
Db 186 YQNAATFG-----ATHSGPLQIAVNOAER-----212
Qy 269 VVFSRNTAVBFDGNVARVGGIYSYGNVAFNLNGKTLFLNNVSPVYIAAKQPTSGQAS 328
Db 213 ---PAQNTAK--NGS---GGALYSDGIDIDQNAVYLFREALTTAI-----252
Qy 329 NTSNNGYDGGAI FCKNGAQAGNNSGS-----VSPDGEVVFSSNVAAKGGAIIYAKK 382
Db 253 -----GKGAVCC-----LPTSGSSTPPVPIVTFSDNKQLVPERNHSIMGGGAIYARK 299
Qy 383 LSVANGCPVQFLRNIA-----NDGAIYLGSGELSLSDYGDIIIFDGNLAKRTAKENAAD 437
Db 300 LSISGGFTLFINNIYSANSQNLGAIAIDTGGELISAERGTTITFQGN-----RTSLPF 354
Qy 438 VNGVTVSSQAISMGSGGKITTLRAKAGHQIILFNDPIEMANGNNOQAQSSKLLKIN---DG 494
Db 355 LNGIHLQNA-----KFLKQIARNGVSIREFYDPI-----TSEADGSTQLNINGDPKN 401
Qy 495 EGYTDIIVFA-----NGSSTLYQNVITIEQGRIVLRKAKLSVNSLSQTGGS--LYM 543
Db 402 KEYTGITLFSGEKSLANDPRFKSTIPQNVNLSAGYLVIKEGAEVTSKFTQSPGSHLV 461
Qy 544 EAGSTWDPVTPQPPQPAANQLTSLNLHLSSLANNVANNVTPPTPPAQDSDHPAIVG 603
Db 462 DLGT-----KLASKEDIAITGLAIDIISL-----SSSSTAIVIK 496
Qy 604 STTAG---SVTISGPPIF-----PEDLDDTAYDRYDMLGNSQKNVLKQLGTGTPPANAP 654
Db 497 ANTANKQISVTDISIELISPTGNAYEDLR-----NRNSQTPELLSLB---PGAGGS 543
Qy 655 SDLTLGNEM---PKGYGQSWKLAWDPTNANGPYTLKATWTKYGNPGPERVASLVPNS 711
Db 544 VVTVAGDFLPVSPHYGPOGNKLAWTGTGNKVGEF-----PNDKINYKPRPEKEGNLVENI 599
Qy 712 LWSGTLDIRSAHSAIOASVDGRS YCRGLWVSGVGNFFVYHDDRDLGQGVYIISGGYSLGAN 771
Db 600 LWGNAVDVRS LMQVQETHASLSQTDRLGLWIDIGNFFHVSASEDNIRYHNSGGTVLSVN 659
Qy 772 SYFGSSMP-GLAFTVFGRSKDYVYCRSNHACIGSVYLSYLTQQALCGSYLF-----GDA 824
Db 660 NEITPKHTYSMAFSLQFSDRDXAVSNNEYR-----MYL-----GSYLYQYTTSLGNI 707
Qy 825 FIRAS-----YFGNQHMKTSYTFEESDVRWDDNCLA 857
Db 708 FRYASRPNVNVGILSRFLQNLPLMI FHLFCAYGHATMDKTDYANFPMVNSWRNNCWA 767
Qy 858 GEIGAGLPVITTPSKYLNELRPPVQAEFSYADHESFTEEGDQARAPKSGHLLNLSVPVG 917
Db 768 IECGSMFLPVENGRLPQGAIPFMKQLVYAYQDFFKETTADGRRRNSGSLTISVPLG 827
Qy 918 VKPRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAPLARGHVGVVRGSM 977
Db 828 IRFEKALXQDVLVDFSPSYIPDFIKRDPSCBAALVISGDSMLVPAAHVSRHAFVGSCTG 887
Qy 978 YASITSNIEVYGHGRYEYRDASRGYGLSAGSRVRF 1012
Db 888 RYHFNDYTELLCRGSI ECRPHARNYNINCGRKFRF 922
```

RESULT 26

```
E72130
polymorphic membrane protein G family CP0761 [imported] - Chlamydophila pneumoniae (stra
Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: E72130; G81541
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A>Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:99206606; PMID:10192388
A/Accession: E72130
A/Molecule type: DNA
A/Residues: 1-841 <ARN>
A/Cross-references: UNIPROT:Q923A1; UNIPARC:UPI00000470B5; GB:AE001586; GB:AE001363; NID
A/Experimental source: strain CWL029
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: G81541
A/Molecule type: DNA
A/Residues: 1-841 <REA>
A/Cross-references: UNIPARC:UPI00000470B5; GB:AE002235; GB:AE002161; NID:g7189672; PIDN:
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: pmp_2; CP0761
C/Supfamily: Chlamydophila pneumoniae polymorphic outer membrane protein G
```

```
Query Match 17.8%; Score 939.5; DB 2; Length 841;
Best Local Similarity 27.0%; Pred. No. 3.6e-49;
Matches 289; Conservative 134; Mismatches 358; Indels 291; Gaps 30;
Qy 1 MQTSFHFPPFLSMILAYSCCSINGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTVTF 60
Db 1 MKIPLRLLISLVPTLSMSNLGAATTEELASNS-FDGTTSSTTSFSSKTSATDGTNYV 59
```

```
Qy 61 SAGELTKNLNDNSTAALPLSCFQ--LLGSFTVLGRGHSITPENI-RTSNGAALSNSAA 117
Db 60 FKDSVVIENVPKTGETOSTCFKNDAAAGDLNLFGLGGFSFTFSNIDATTASGAAGSEAA 119
Qy 118 DGLPTIEGFKELSPFNCNLSLLAVLPAATTNKSGTPTTSTPNSNGTIYSKTDLLANNEK 177
Db 120 NKTVTLSGFALSF-----LKSFASTVINGLGAIVNKNLSLLDNDK 161
Qy 178 FSPYSLNVSGDGGAI--AKSLTVQGISKLCVFOBNTAQADGACQVVTFSAMANEAPIA 236
Db 162 VLIQDNFSTGCGGAINCAGSLKI-----ANNKSL 191
Qy 237 FVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVVPSFRNTAVBFDGNVARVGGIYSYN 296
Db 192 FIGNSSSTRGGAHT-----KN 208
Qy 297 VAFNLNCKTLFLNNVSPVYIAAKQPTSGQASNTSNNGYDGGAI FCKNGAQAGNNSGSV 356
Db 209 LTLSSGGETLFGNTA-----PT-----226
Qy 357 SPDGEVVFSSNVAGKGAIYAKKLSVANCGVPQFLRNIA DGGAIYLGSGELSLSA 416
Db 227 -----AAGK-----GGAIAADSGTSLISG 246
Qy 417 DYGDIIIPDGNLKR TAKENAADVNGVTVSSQAISMGSGGKITTLRAKAGHQIILFNDPIEMA 476
Db 247 DSGDIIIFEG-----NTIGATG-TVSHSAIDLGTSAKITALRAAAGHTIYFYDPTIV- 296
Qy 477 NGNNQPAQSSKLLKIN-----DGEYGTGDIIVFA-----NGSSTLYQNVITIQ 518
Db 297 TGSTSVADA---LNINSPDTGDNKEYTGTIVFSGEKLTEAEKDEKRTSKLLQNVAFKN 353
Qy 519 GRIVLREKAKLSVNSLSQTGGS--LYMBAGSTWDPVTPQPPQPAANQLITLSNLHLSLS 577
Db 354 GTVVLKGDVVLSANGPFSQDANSKLINDLGTSL-----VANTESIELTWLEINID 402
Qy 578 SLLANNVTPPTPPAQDSDHPAIVGSTTAGSVTISGPIFFEDLDDTAYDR-----Y 629
```



C:Genetics:

A:Gene: pmpI

C:Superfamily: Chlamydo

phila pneumoniae polymorphic outer membrane protein G

Query Match 14.6%; Score 771.5; DB 2; Length 878;  
Best Local Similarity 26.5%; Pred. No. 6.3e-39;  
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

QY 53 DPSGTTVTSAGELTKNLDNSIAALPLSCFQNL-----GSTVLGRGHSHTPE 101  
DB 26 DPLGETA-----LTKNPNHVCTFECDCTMESLPALCAHSDQDPLVLGNSCMFVS 80  
QY 102 NIRSTNGAALSNSAADLFTIEGKELSPNCNLSLLAVLPAATNKGSGQTPTTSTPSN 161  
DB 81 KLHITDPEALFKEGD--LSIQNPRFLSFIDCSS-----KESSPS- 119  
QY 162 GTIYSKT-DLLLLNNEKSFYSNLVSGDGGDAIDAKSLTVQGISKLCVQENTAOADGAC 220  
DB 120 -IIHKNGQLSLRNGNSPCRNHABGGGAIADAPSLQHNLYLFTAFENSSKNGGAI 178  
QY 221 QVTSFSAMANEAPIAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNVAVEF 280  
DB 179 QAQT-FSLRSVPISFARNADLNGGAIC----- 208  
QY 281 DGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVVIAAKQPTSGQASNTSNNYDGGAI 340  
DB 209 -----SNLICGNV-----NPLFF-----TGNATNGGAI 233  
QY 341 FCKNGAAGSNNSVSDGEGVVPFSSNVAAAGKGAIYAKKLSVANGPVQFLRNIAAD 400  
DB 234 CCI--SDLNTSEKGSLSLACQETLFASSAKGKAIYAKHVLRYNGPVFINNSAKI 291  
QY 401 GGAIYLGSGHLSLADYDIIIDGNLKRKTAKENADVNGVTSQAISMGGKITTLLR 460  
DB 292 GGATAIGGSGLSILAGSGVLQNNQRTSQGLVR-NAIYLRKDAI-----LSSLE 343  
QY 461 AKAGHQILFNPI-EMANGNQPAOSSKLLKINDGEYTG-----DIVPA----- 504  
DB 344 ARNG-DILFFPPIVQSSSKSPSSLOASVTSPTPATASPLVITQTSANRSVFSRRL 402  
QY 505 -----NGSSTLYQNTYIQGRVLRKAKLSVNSLSQTGGS-LYNEAGSTWDFVTPQ 555  
DB 403 SEBEKTPDNLTSQLOQPTELKSGRLVKDRAVLAPSLSQDPQALLIMEAGTS----- 455  
QY 556 PPOQPPAANQLITNLHLSLSLLANNAVTPNPPAQAQSHPAVIGSTTAGSVTISGP 615  
DB 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP 485  
QY 616 -----IPFEDL-DDTAYDRYDMLGSKNINVLKLGTKPPANAPSDLTL--GNEMPKY 666  
DB 486 NLSIQKIFLSNGDENFYENVLLSKQN-NIPLTL-----SKEQSHLHLPDGNLSHP 539  
QY 667 QYQSWKLANDPNTANNPYTLKATWTGTGYNPQPERVASLPNSLWGSILDIRSAHSAI 726  
DB 540 QYQGDWTFM--KDSDEG-HSLIANWTPKNVPHPERQSTLVANLWNTYSMDQAVQSMI 596  
QY 727 QASVDGRYCGLWVGVSNPFY-HDRDALG-QGYRYISGGYSIGANSY-FCSSMFLGAF 783  
DB 597 NTIAHGAGYLFGTWGSAYSNLFYAHDSGKPIDNWHRSGLYLGISHTLDDSHFCIAA 656  
QY 784 TEVFGSKDYVVCRSNHHACIGSVYLSQALCGSYLFGDAPIASYCFGNQHKTSY-T 842  
DB 657 GOLLGKSDSPITSTE-----TTSYIATVQQLATPLM-KISAQACYNESITHELTKYRS 710  
QY 843 FAEBSDVWRDNNCLAGIEGAGLPIVITPSKLYLNLPRFPVQAFPSYADHESFTEBGOAR 902  
DB 711 FSKGFGSMHSAVSGEVCSAIPVNSGSLF-SFPSIFSKLQGFSGTQDGFESSGEIR 769  
QY 903 APKSGHLLNLVPGVKFDRCSSTHPNKYSFMAAYICDARTISGTETTLTSHQETWTD 962  
DB 770 SFSASSFRNISLPMGITPEKKSQKTRNTYTFILGAYIQDLKRDVBSGPPVLLKNAVMDAP 829  
QY 963 APLARHGVVVGSMYASLITSNIEYVGHGREYDRDASRGYGLSAGSRVRF 1012

Db

330 MANLDSRAYPRLTNQAL-HRQTLNVSIVLGGQSHSYSLDGLTTRF 878

RESULT 29

E86492

polymorphic outer membrane protein G family [imported] - Chlamydo

phila pneumoniae, Chlamydia pneumoniae

C:Species: Chlamydo

phila pneumoniae, Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002

C:Accession: E86492

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349; PMID:10871362

A:Accession: E86492

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-712 &lt;STO&gt;

A:Cross-references: UNIPARC:UPI0000165697; GB:BA000008; NID:g8978386; PIDN:BA098223.1; G:

C:Experimental source: strain J138

A:Gene: pmp 2.1

C:Superfamily: Chlamydo

phila pneumoniae polymorphic outer membrane protein G

Query Match 13.3%; Score 703; DB 2; Length 712;

Best Local Similarity 25.1%; Pred. No. 6.7e-35;

Matches 246; Conservative 109; Mismatches 304; Indels 320; Gaps 33;

QY 1 MQTSFHKFFLSMILAYSCCSLNGGYYAAEIMVPGQIYDGETLTVSFPVTVIGDPSGTTVP 60

DB 1 MKIPLRLLLSLVPFLSMLNLLGAATTEELSASNS-PDGTTSSTSPSKTSSATDGTNYV 59

QY 61 SAGELTKNLDNSIAALPLSCFQ--LLGSFTVLGRGHSHTPENI-RSTNGAALSNSAA 117

DB 60 FKDSVWIEVPKTGETQSTCFKNDAAAGDLNPLGGGFSPTFSNIDATTASGAIGSEAA 119

QY 118 DGLFTIISGPELSPNCNLSLLAVLPAATNKGSGQTPTTTPSNGTITSKTDLLIANNEK 177

DB 120 NKTVTLSGFSALSF-----LKSPASTVTNGLGAINVKGNSLLDNDK 161

QY 178 PSFYSNLVSGDGGDAID-AKSLTVQGISKLCVFQENTAQADGACQVTVSFSAMANEAPIA 236

DB 162 VLIQDNFSTGCGAINCAGSLKI-----ANNKLS 191

QY 237 FVANAVGVRGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVPEPDGNVARGGIIYSGN 296

DB 192 FIGNSSSTRGAIHT-----KN 208

QY 297 VAFLNNGKTLFLNNVASPVVIAAKQPTSGQASNTSNYDGGAIFCNGAQAQSNNSGV 356

DB 209 LTLSSGGETLFGGNTA-----PT----- 226

QY 357 SFDGEGVVPFSSNVAAAGKGAIYAKKLSVANGCPVQFLRNIAINDGGAIYLGESGELSLSA 416

DB 227 -----AACK-----GGAIAIADSGTSLISG 246

QY 417 DYGDIIIPDGLKRTAKENAAADVNGVTVSSQAISMGGKITTLLAKAGHQILFNDPIEMA 476

DB 247 DSGDIIPEG-----NTIGATG-TVSHSAIDLGTSAKITAKRAAQCHTIYFYDPTIV- 296

QY 477 NGNNQPAOSSKLLKIN-----DGEYTGDIVPA-----NGSSTLYQNVTIEQ 518

DB 297 TGSTSVADA---LWNSPDTGDNKEYTGTIVFSGEKLTEAEKDEKRTSKLLQNVAFKN 353

QY 519 GRIVLRKAKLSVNSLSQTGGS-LYMEAGSTWDPVTPQPPQPPAANQLITSLNHLISLS 577

DB 354 GTVVVLKGDVLSANGFSQDANSKLMDLGTSL-----VANTESLTWLEINID 402

QY 578 SLLANNAVTPNPPAQAQSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDR-----Y 629

DB 403 SLRNGKKI-----KLSAATAQKDIRIDRPVVLAISDESFYQNGFLNEDHSY 448

QY 630 DMLGSKNINVLKLGTKPPANAPSDLTLGNEMPKYQYQGSWKLANDPNTANNPYTLK 689

Db	449 D-----GILELDAGKDIVISADSRSDAVQSP-YGVQKWTINWSTDDK-----K 492	
Qy	690 AT--WTKGYNPGERVASLVNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGVSNF 747	
Db	493 ATVSWAKQSFNTABQEAFLVNLWGSFIDVRSPQNFTELGTGAPYEKRFVWAGISNV 552	
Qy	748 FPHDRDALQGVYRISGGYSLGANSYP-GSSMFGLAFTVEVGRSKDYVVCNRRHACIGS 806	
Db	553 LHRSGRENQRKFRHVSGGAVGASTRMPCGGDTLSLGFALFARDKDYFANTNTFAKTYAGS 612	
Qy	807 VYLSQQALCGSYLFGDAFIRASVYFGNQHMKTSYTFABESDVRWNNCLAGEIGAGLPI 866	
Db	613 LRL-----QHDASLYSVV-----SILLGE--GGLRE 636	
Qy	867 VTPSKLYLNELRP---FVQAEFSVADHESFTEE-----GDQARAFKSGHL 909	
Db	637 ILLP---YVSKTLPSCFYQLSYGHTDRMKTESLPPPPRRRRRIILLGEDMSGLSEWEL 693	
Qy	910 -LNLSPVGVKFDRCSSTH 927	
Db	694 ELLKIPAAEDFSK-STLH 711	
RESULT 30		
G72076		
polymorphic outer membrane protein h family - Chlamydothila pneumoniae (strain CWL029)		
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae		
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004		
C:Accession: G72076		
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, J.;		
Nature Genet. 21, 385-389, 1999		
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.		
A:Reference number: A72000; MUID:99206606; PMID:10192388		
A:Accession: G72076		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-978 <ARN>		
A:Cross-references: UNIPROT:Q92895; UNIPARC:UPI0000046ABB; GB:AE001629; GB:AE001363; NID		
A:Experimental source: strain CWL029		
C:Genetics:		
A:Gene: pmp.14		
C:Superfamily: Polymorphic membrane protein H family		
Query Match 13.2%; Score 693.5; DB 2; Length 978;		
Best Local Similarity 25.6%; Pred. No. 4e-34;		
Matches 278; Conservative 147; Mismatches 412; Indels 251; Gaps 42;		
Qy	45 SPPTVIGDPSGTTVFSAAGELTKNLD---NSIAALPLSCFGLNLAGSFTVLGRGSLTF 100	
Db	22 AFAETRLGGNFVPPITNQGEIILLTDFVCSNPLGASFSFSSFNSSNLSLLKGLSLTF 81	
Qy	101 ENIRTSNGAALSNSAAGLFTIEGFKLSFSCNLSLAVLPAATTKGSGTPTTSTPS 160	
Db	82 TSCOAPTSNYALLSAETL-TFKNSSINF-----GNQ-----STGL 119	
Qy	161 NGTIYKTDLLLNNEKFSYSLVSGDGAIDAKSLTVGSKLVCFQONTQAQAGGAC 220	
Db	120 GGLIYCK-DIVFQSIKDLIFTTNRV---AYSPASVTSATPAI-----TTVTTCASAL 168	
Qy	221 QVTSFSAMANEAPIAVANVAVRGGIAAVQDGGQGVSSSTSTEDPVSFSRNTA-VE 279	
Db	169 QPTDSLTVENISQISKFPGLNLF-----SAISSSPFVAVVKFINNTATMS 214	
Qy	280 FDGNVARVGGG-IYSYGNVAFPLNKTFLNNVASPVYIAAKOPTSQASNTSNVYDGG 338	
Db	215 FSHNFTSSGGVYVGGSLFLFERNSSGCIIFTANSCVNSLKGVPSSGTYA-----LGSGG 269	
Qy	339 AIFCKNGAAGSNNSGVSFED-----GEGVFFSSNVAAKGKGA 378	
Db	270 AICPTGTPELKNQKCTFSYNGTPNDAGIAVETCNVGNQALLDSDNTAARNGAI 329	
Qy	379 YAKKLSVANGCPVQFLRNIANDGAIYLG-----ESGELSLSADYDGIILFDGNLKR 430	

Db	330 CAKVLNIQGRPIBFSRRNAEKGGAIFGPSVGPAPKQKTSTLTILASEGDIAFQGNMLNT 389	
Qy	431 AKENADVNGVTVSSQAISSMGSKITTLRAKAGHOILLFNDPIEMANGNNOPAOSSKLLK 490	
Db	390 KP-----GT---RNAITVAGGEIVLSLQAGSRIVFYDPIYHSLFTTSP--SKNDIT 437	
Qy	491 INDGEGYTGDIYF-----ANGSSTLYQNVTIEQGRIVLRKAKLSVNSLSQT 537	
Db	438 IN-ANGASGVVFTSKGLSSTELLIPANTTTILLGTVKIASGELKITDNVAVVNLGFATQ 496	
Qy	538 G-GSLYMEAGSTWFTVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPPTPPAQD 596	
Db	497 GSGQLTLGSGGTGLGAT---PTGAPAAVD-FTIGKLAFPDPSFLKRDVFS----- 542	
Qy	597 SHPAVIGSTTAG--SVTISGPIFPEDLDDTAYDRYDMLGNSQKINVLKQLGCTKPPANAP 654	
Db	543 -----ASVNAGTKNVTUTGALVLDEHVT--DLYD-----MWSLQSPVAIP 581	
Qy	655 SDLTILGNMPK-----YGYQGSWKLAWD-----PNTANNNGPYTLK 689	
Db	582 LAVFKGATVTKGPDGGEIATPSSHVGYQGWSTWRSRPLLIAPDGPFGPSPSANTLY 641	
Qy	690 ATW-----TKGYNPGERVASLVNSLWGSILDIRSAHSAIQ--ASVDGRSYCRGLWYS 742	
Db	642 AVWNSDTLVRSTYILLDPERYGEIVNSLSLWISFLGNQAFSDIILQDVLIDH----PGLSIT 697	
Qy	743 GVSNEFFYHDDRDLG-----QGVYISGGV-----SLGANSYFGSSMFLGAPTEVFOR 789	
Db	698 A-----KALGAYVHTPPQGHGFSRGYGGYQAALSMN-YTDHTLGLSFGQLYOK 747	
Qy	790 SKDYVVCNRRHACIGSVYLSVTLQALCGSYLFG-----DAFI--RASYGFGNQHMK 838	
Db	748 TN-----ANPYDSRCSEQMYLLS-----PFGQFPVITQKSEALISWKAAYGYSKNHNL 795	
Qy	839 TSY---TPAESDVRWNNCLAGETGAGLPI-----VTPSKLYLNELRPFVQAEFSYADH 891	
Db	796 TTYLRPDKAPKSGQGHNSYVVLISAEHPFLNWCILTRPLAQAWDLDSGFIASFLGGWQ 855	
Qy	892 ESFTREGDQARAFKSHLLNLSPVGVKFDRCSS-----THPKNYSFMAAYICDAYT 944	
Db	856 SKFTETGDLQRSFSRGKGYNSLPIG-----CSSQWTFPKKAPSTLTITKLAYKPDYIRV 910	
Qy	945 ISGTETTLSSHQETWTTDAPHLARHGVVVRGSMYASLTSNIEVYGHGREYEDASRGYGL 1004	
Db	911 NPHNIVTVVNSQESISGANLRHGLFVQIHVDVLDLTEDTQAFNLNYTFDGKNGFTNHRV 970	
Qy	1005 SAGSRVRF 1012	
Db	971 STGLKSTF 978	
RESULT 31		
C86547		
polymorphic outer membrane protein H family [imported] - Chlamydothila pneumoniae (stra		
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae		
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004		
C:Accession: C86547		
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I		
Nucleic Acids Res. 28, 2311-2314, 2000		
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.		
A:Reference number: A86491; MUID:20330349; PMID:10871362		
A:Accession: C86547		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-978 <STO>		
A:Cross-references: UNIPROT:Q92895; UNIPARC:UPI0000131C83; GB:BA000008; NID:98978825; P		
A:Experimental source: strain J138		
C:Genetics:		
A:Gene: pmp.14		
C:Superfamily: Polymorphic membrane protein H family		
Query Match 13.1%; Score 688.5; DB 2; Length 978;		



Db 543 -----ASVNAKTVLTCALVLEHDT--DLVD-----MVSQSPVAIP 581  
Qy 655 SDLTLGNMPK-----XYQGSWKAWD-----PNTANNQPYTLK 689  
Db 582 IAVFKGATVTKTGPDEIATPSHYGQKWSYTWSPILLIPAPDGGPFGSPSANTLY 641  
Qy 690 ATW-----TKTGNPFPERVASLVNSLWGLDIRSAHSAIQ--ASVDGRSYCRLAWS 742  
Db 642 AVNSTDVLRSTYILDPERYGVIVNSLWISFLGNQAFSDIILQDVLIDH-----PGLSIT 697  
Qy 743 GVSNPFPYHDDALG-----QGVYRISGGY-----SLGANSYFGSSMFLAFTVEFGR 789  
Db 698 A-----KALGAYEHTPQGHGFGSGRYGGYQAALSMN-YTDHTTLLGLSFGQLYK 747  
Qy 790 SKDYVYVCRSHHACIGSVYLSQOALCGSYLFG-----DAFI--RASGFGNQHKM 838  
Db 748 TN-----ANPYDRSCSQMVLSS-----PFGQPPVTVQKSEALISWKAAYGYSKHLN 795  
Qy 839 TSY-----TPABESDVRWNNCLAGEICAGLPI-----VITPSKLYLNELRFPVQAEFSYADH 891  
Db 796 TTYLRDPKAPKQCGQWHNSYYVLLISAEPFLNWCILLTRPLAQAWDLGFGISAFLGGWQ 855  
Qy 892 ESTEEDQBARPKSGHLNLSVPVGVKFDRCSS-----THPNKYSFMAAYICDAYRT 944  
Db 856 SKFTETGDLQSPSGRGKYNVSLPIG-----CSSQWFTPKKAPSTLTIKLAYKPDYRV 910  
Qy 945 ISGTETTLASHQSTTTDAPHLARHGVPVGRVMSYASLTSNIEVYGHGREYERDASRGYL 1004  
Db 911 NPHNIVTVSNQDSTISGANLRHGLFVQIHVDVLDLTEDTQAPLNTYFDGKNGFTNHRV 970  
Qy 1005 SAGSRVRP 1012  
Db 971 STGLKSTP 978

RESULT 33  
F81722  
polymorphic membrane protein E/F family TC0261 [imported] - Chlamydia muridarum (strain C); Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C; Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C; Accession: F81722  
R; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A; Reference number: A81500; MUID:20150255; PMID:10684935  
A; Accession: F81722  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-976 <TET>  
A; Cross-references: UNIPROT:Q9PL47; UNIPARC:UPI0000057865; GB:AE002293; GB:AE002160; NID  
A; Experimental source: strain Nigg (MoPn)  
C; Genetics:  
A; Gene: TC0261

Query Match 12.9%; Score 679.5; DB 2; Length 976;  
Best Local Similarity 26.1%; Pred. No. 2.9e-33;  
Matches 265; Conservative 150; Mismatches 413; Indels 187; Gaps 38;

Qy 78 PLSCFGLLGSTVLRGHSIT-----FENI-----RTSTNGAA-----LSNSAAD 118  
Db 33 PYHMTGLFPKPKVNLGDTNLTLDYHLNLCILACLOQTPYEGAAFTVDTYLSFSDTKD 92  
Qy 119 GLPTIEGKELSPNCNSLLAVLPAATNTKSGSQPTTSTPSNGTIYSKTDLLLNNEKF 178  
Db 93 GIFC-----PKNLT-----PESGGVIGSPQNTPTIKIHNTI-----GPVLFNNTCH 135  
Qy 179 SFYSNL-----VSGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVT----- 224  
Db 136 RLMTQTDPENEGNKAREGALHAGDVIYSNNQNLVGFINKFYVQGGAISANTPAYKENK 195  
Qy 225 -SPSAMANEAPIAFVNAVGRGGGIAAVQDQGGQSVSSSTSTEDPVPVFSRNTAVEFDGN 283

Db 196 SSFLCLANSC-----IQTKTGKGGAIY-----VSTSCSFE-----NNKDLFLFTQN 237  
Qy 284 VARVCGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSCQASNTSNNTYDGDGAIFCK 343  
Db 238 SGCAGGAIPTSCSLIGNQGDIVFYSNGFK-----NVDNATNESDGGAI--K 284  
Qy 344 NGAQAGSNNSVSPDGEVGVFFSSNVAAGKGGAIYAKKLSVANGCPVOFLNIN--DGG 402  
Db 285 VTTRLDITNGS-----QIFFSDNISRNFGGAIHAPCLHLVGNPTVFTNNIANHTGG 337  
Qy 403 AIYLGESGELSADYGDIIIDGNLKRKAKENAADVNGVTVSS-----QAISMGSGGKIT 457  
Db 338 AIYITGTETSKISADHAIIFDNNI-----SANATNADGSSNTNPHRNAITWDSAGGI 393  
Qy 458 TLKAKAGHQILPNDPIEMANGN-----NOPAQSSKLLKINDGEGTGDIVFANGSSTLYQ 512  
Db 394 ELGAKSQNLPIPYDPIQVNTAGVTDFNKDASQTCVVFSGATVLSADISQANLQTKTPA 453  
Qy 513 NVTIEQGRIVLRKAKLSVNSLSQTGSLYMEAG---STWDFVTTPQPQPPP---AANQL 566  
Db 454 TLTSLHGLLCIEDRAQLTVNNFTQTGGIVALGNGAVLSSYQHSITTDATQTPTTTTDDAS 513  
Qy 567 ITLSNLHLSSLANNNAVTPNPP--AQSHPAVIGSTTAGSVTISGPI--FPEDLD 622  
Db 514 VTLNHLGLNLPILKOGAEMPLLWVEPISTTCQNTTYYTSDTAASFSLNGATLSLDEDG 573  
Qy 623 DTAYDRYDLGWSNQKINVLKOLGTFP--PANAPSDLTIGN-----MPKYGYQGSW 672  
Db 574 NSPYENTD-----LSRALYQPMALASEANDQLOESMDPFSKVNPHYGMQGLW 623  
Qy 673 KLAW-----DPNTANNQPYTLKATTKTCYNPQPERVASLVNSLWGSIL 717  
Db 624 TWGNAKTENPTTTPPATITDPKKAQFHTLLTLWLPAGYIPSPKHSPLIANTLWGNIL 683  
Qy 718 DIRS--AHSIAQASVGRSYCRGLWVGSVSNFPYHDDRALDQGYRVIISGYSLG---ANS 772  
Db 684 FATENLKNSSQGLLDRPFW--GITGGGLGMVYQEPKDHGPHGHTSGYSAGMITGNT 741  
Qy 773 YFGSSMFLAFTVEFCR--SKDYVYVCRSHHACIGSVYLSQOALCGSYLFGDAFIRASYG 831  
Db 742 HTPSLRFSGSYTKLNEYAKNYV--SSKNYSQGEMLLSLQEGMLMTKLIG-----LYS 793  
Qy 832 FGNQHKMSTVYTAEB--SDVRWNNCLAGEICAGLPI-----VITPSKLYLNELRPFV 882  
Db 794 YGNHNSHFTTQGEDLSSQGEFHSQTFEGAVFDELKPKPGRTHILT-----APPL 844  
Qy 883 QAEFSYADHESPTTEGDQAPAP--KSHLLNLSVPVGVKFDRCGSTH--PNKYSFMAAYICD 940  
Db 845 GAIGMYSKLSSPTEVGAYPRFTFITETPLINVLPIGVKGSFMNATHRPQAWTVLAYQPV 904  
Qy 941 AYKTIISGTETTLSSHQETWTTDAPHLARHGVPVGRVMSYASLT--SNIEVYGHGRY 993  
Db 905 LYRQEPSISTQLLAGKGMWFGHGPSASRALAYKISQKTQLLRFATLQYHGY 959

RESULT 34  
E71460  
probable outer membrane protein E - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C; Species: Chlamydia trachomatis  
C; Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
C; Accession: E71460  
R; Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell Science 282, 754-759, 1998  
A; Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra A; Reference number: A71570; MUID:99000809; PMID:9784136  
A; Accession: E71460  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-964 <ARN>  
A; Cross-references: UNIPROT:O84877; UNIPARC:UPI00000131CF3; GB:AE001360; GB:AE001273; NI  
A; Experimental source: serotype D, strain UW-3/Cx  
C; Genetics:

A:Gene: pmp2

```
Query Match      12.3%; Score 649.5; DB 2; Length 964;
Best Local Similarity 25.8%; Pred. No. 1.9e-31;
Matches 260; Conservative 155; Mismatches 389; Indels 205; Gaps 43;

Qy 84 NLLGSPTVLGRHSILT--PENIR-----TSTNGAALSNSAADGLTIEGPKELSPS 132
Db 39 SLNKSISUTGTHNUTCNVCLNRYLAILLOKTPNEGAIV--TITDYLSPFDTOKEGIYF 96

Qy 133 NCNLSLAVLPAATTKGSGTPTTSTPBGNGTIYSKTDLLLNNE--KFSFYSLNLSVGD-- 188
Db 97 AKN-----LTPESGGAIGVSPNSPTVEIRDIT---GPVIFENNTCCRLFTWRNPVAADKI 149

Qy 189 --GGAIKSLTVQGISKLCVQFQNTAAGGACQVTVSPSAMANEAPIAFVAVV----- 241
Db 150 REGGAIAHQNLVYINHHVDVGMKNSFYVQGAISTANTFVVSQSCFLPMDNICIQTN 209

Qy 242 -AGVGGGIAAVQDQCGQSVSSTSTEDPVVSPSRNTAVEFDGNVARVGGGIYS----- 293
Db 210 TAG-KGGAIYA-----GTSNSFE-----SNCDLFFINNACCAGAIKSPICSLTG 254

Qy 294 -YGNVAFLLNGKTLFLNNVASPVYIAKQPTSGQASNTSNNGDGAIFCKNGAQAGSNN 352
Db 255 NRGNIYFVNN--RCPKN-----VETASSEASDGAIKVTRLDVTGN- 294

Qy 353 SGSVFDGEGVVFPSSNVAAGKGGAIYAKLSVANCQVTVSPSAMANEAPIAFVAVV----- 411
Db 295 -----RGRIFPFDNITKNYGGAIYAPVTVLDNGPTYFINNIANNKGAIVIDGTSN 346

Qy 412 LLSADYDGLIPDGNLKRKTAKENAADVNGVTSS-----QALSMGSGKITTIRAKAGHQ 466
Db 347 SKISADRAHIIIFNENI-----VTNVTNANGTSTSAMPPRRNAITVASSGEILLGSSQN 402

Qy 467 ILFNDPIEMANGNNQPAOSSKLLKINDGEGYTGDIVFAN--GSGSTLYQN-----V 514
Db 403 LIFDPIEVSN-----AGVSVSFNKEADQTSVVFSGATVNSADPQRNLQTKTPAPL 455

Qy 515 TIBQRIVLRLKAKLSVNSLSQTSGLSYMEAGS-----TWDPVTPPQQPPAANQL 566
Db 456 TLSNGFLCIEBHAQLTVNRFTQTGGVWSLNGAVLSCKYKNGTGD-----SASNAS 505

Qy 567 ITLNLHLSSLILLANNA-----VTNPPTPNPAQDSHPAIVGST-TAGSVTISGPIFPEDL 621
Db 506 ITLHGIGLNTLSILSGARIEPLLWVEPTN--NSNNYTADTAATFSLSDVKLS-----LIDDY 560

Qy 622 DDTAYDRYDMLGSKINVLKQLGQTKP--PANAPSDLTLDGNE-----MPKYGYQGS 671
Db 561 GNSPYESTD-----LTHALSSQPMLSISEASDNLQLOSENIDFSGLVNPHYGMQGL 610

Qy 672 WKLAW-----DPNTANNQPYTLKATWTGTGYNPQGPVAVSLVPSNLSMGS- 715
Db 611 WTWGAKTQDPEPASSATITDPKANRRPHRTLLLTWLPAGVYVPP8PKHRSPLIANTLMGNM 670

Qy 716 ILDIRSAHSAIQASVDGRSVCRLMVSGVSNFFYHDDRDLALQGYRIYISGGYSLG--ANS 772
Db 671 LLATESLKNSEALTPSGHPFW-GITGGGLGMVWQDPRNHPGPHMRSGSGYSAGMIAGQT 729

Qy 773 YFGSSMFLGAFTEVPGR-SKDYVVCRSNHHCIGSVYLSLTOQA-----LCGSYLFQDAPI 826
Db 730 HTPSLKFSQTYTKLNERAKNNV--SSKNYSQCGEMLFSLQEGFLLTKVLGLYSYGDHNC 787

Qy 827 RASVFGNQHKMTSYTPABESDVWNNCLAGEIGAGLPVITPYSKLYNELRPFVQABF 886
Db 788 HHFTYQG-ENLTSQGTFRASQT-----MGAFFVFDLPMPKFPFGSTHILT--APFLGALG 836

Qy 887 SYADHESPTREGDQARAFKS-CHILNLVSPVGVKFDRCSSTH-PNKYSFMAAYICDAVRT 944
Db 837 IYSSLSHFTEVAGYPRSFSTKTPNLINLVPIGVKGSFNNATHRPOQMTVELAYOPVLYRQ 896

Qy 945 ISGTETTLSSHQETWTDDAPHLARHGVVVRGSMYAS-----LTSNIEVYG 989
Db 897 EPGIAAQLASKGIWFGSGSPSSRHAMSYSKISQQTQPLSWLTLHPQYHG 945
```

RESULT 35  
F86548

polymorphic outer membrane protein E family [imported] - Chlamydothila pneumoniae (strain C) Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: F86548  
R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ishi, Nucleic Acids Res. 28, 2311-2314, 2000  
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
A/Reference number: A86491; MUID:20330349; PMID:10871362  
A/Accession: F86548  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-938 <STO>  
A/Cross-references: UNIPROT:Q92883; UNIPARC:UPI000004707C; GB:BA000008; NID:G8978836; PI  
A/Experimental source: strain J138  
C/Genetics:  
A:Gene: pmp\_15

Query Match 12.2%; Score 644; DB 2; Length 938;  
Best Local Similarity 25.8%; Pred. No. 3.9e-31;  
Matches 266; Conservative 160; Mismatches 408; Indels 198; Gaps 44;

```
Qy 81 CFCNLLG-SPTVLGRHSLTFENIRTSNTNGAALS---NSAADGLFTIEGPKELSP-SNQN 135
Db 5 CFCNLLPFTFVLANEGQLPLETYIT-----LSPEYQAAPQVGFTNQNDLAIVGNHN 58

Qy 136 SLLAVLPAATTKGSGT-----PTTSTPBGNGTIYSKTDLLLNNEKFSF 180
Db 59 DFILDYKYRSGALCKNLLISENGVFFERKXVCPNSGGAIYAAQNTCISKNQYAF 118

Qy 181 YSNLVSVD-----GGAIKSLTVQGISKLCVQFQNTAAGGACQVTVSPSAMAN 231
Db 119 TTNLVSNDPTATAGSLGGALFALNCSTTNLNGQTFVDNLALNKGALYETNLSIKON 178

Qy 232 EAPIAFVANVA---GVGGGIAAVQDQCGQSVSSTSTEDPVVSPSRNTAVEFDGNVARVG 288
Db 179 KGPTIIIKQNRALNSDSLGGGIYS-----GNSLNI-----GNSGAIQITSNSSGSG 224

Qy 289 GGIYSYGNVAFLLNGKTLFLNNVASPVYIAKQPTSGQASNTSNNGDGG-----AIFCKN 344
Db 225 GGIFSTQTLTISNKKLIEISE-----NSAFANNYGNFNPGGGLTTTTC-- 270

Qy 345 GAQAGSNNSGVSFDGEGVYVFPSSNVAAGKGGAIYAKLSVANCQVTVQFLNIANDGCAI 404
Db 271 -----TILNRRBGLV-FNNQSQSGAIHAKSIKENGPPVFLNNTATRGAL 319

Qy 405 YL-----GESGELSADYGDIIIPDGNLKRKTAKENAAD--VNGVTVSSQAISMGGSGKIT 458
Db 320 LNLGAGSGNSGFLSADNGDIIIFNNN---TASKHALPPPYRNAIHTSPNNMLQI----- 371

Qy 459 LRAKAGHOILLFNDPIEMANGNNQPAOSSKLLKINDGEGYTGDIVFANGSSSTLYQNVT--- 515
Db 372 --ARPGRVLYFDYDIE-----HELPSFPLIFNFETGHTGTVLFS--GEHVHQNFTDEM 421

Qy 516 -----IEQRIVLREKAKLSVNSLSQTSGLSYMEAGS--TWDPVTPPQQPPA 562
Db 422 NFPSYLRNTSELRGVLAVEDGAGLACYKFPQGGTLLLGQGAIVITAGITPSTPTT 481

Qy 563 ANQLITLNLHLSSILLANNAVNTNPTNPPAODSHPAVIGST-----TAGSVTISGPIFP 618
Db 482 VGSTITLNLHAIDLPSILSFOA-----QAPKIWIYPTKGTSTYTEDSNPTITISGTLTL 535

Qy 619 EDL-DDTAYDRYDMLGSKINVLKQL-QLGKTPANAPSDLTLDGNEPKYGYQSGKLA 676
Db 536 RNSNNEPDYDLSLHSLKLEKVPLLYIVDVAQKINSQDLSLUNSGEHTGYQINWTY 595

Qy 677 DPNT-----ANNQPYTLKATWTGTGYNPQGPVAVSLVPSNLSMGSILDIRSAHSAI 726
Db 596 VETTTITNPTSLLGANTKHKLLYANWSPGVYRPHPRRGRFITNALWQSYATLALGHL 655
```



```
QY 727 QASVDGRSVCRLGWGVSNPFYHDDRALG-QGYVYISGGYSLGANSYFGSS-MFGLAFT 784
DB 656 SSWDEKHAASLOGIGL---LVHQDKNGFGRFSGHMTGYSATTATSSQSPNLSLGA 712
QY 785 EVFGRSKDYV--CRSNHACIGSVYLSLQALCGSYLFGDAFIRAS-----YGFQGNQHM 838
DB 713 QPFSKAKEHESQNSSTSSH---YFS---GMCIENTLFKEWIRLSVSLAYMFTSEHTH 763
QY 839 TSYTFAEE--SDVRDNDNCLAGEIGAGLPVITPS-----KLYLNELEPPFQAEFSYAD 890
DB 764 TMYQGLLEGNSQGSFHNHTLAG---ALSCVFLPQPHGESLQIY-----PFITALAIRGN 814
QY 891 HESFTEEGDQARAFKSGH--LLNLSPVGVKDRCSSTHPNKYSFM-----AAVYICDAYRT 944
DB 815 LAAPQESGDHAREP-SLHRPLTDVSLPVGI---RASWKHHRVPLVWLTEISYRSTLYRQ 870
QY 945 ISGTETTLTSHOETWTTDAFLARH--GVVVRGSM--YASLTSNIEVYGHGREYRDASR 1000
DB 871 DPELHSLKLLISQGTWTTQATPVYTNALGIKVKNTMQVPPKVTLSLDYSA---DISSSTL 926
QY 1001 GYGLSAGSRVRF 1012
DB 927 SHYLNVASRMRF 938

RESULT 36
H72074
polymorphic membrane protein E/F family CP0286 [imported] - Chlamydothila pneumoniae (sp
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: H72074; E81593
R:Kalmann, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: H72074
A:Molecule type: DNA
A:Residues: 1-938 <ARN>
A:Cross-references: UNIPROT:Q92893; UNIPARC:UPI000004707C; GB:AE001631; GB:AE001363; NID
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: E81593
A:Molecule type: DNA
A:Residues: 1-938 <REA>
A:Cross-references: UNIPARC:UPI000004707C; GB:AE002190; GB:AE002161; NID:g7189209; PIDN:
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: pmp_15; CP0286

Query Match 12.2%; Score 644; DB 2; Length 938;
Best Local Similarity 25.8%; Pred. No. 3.9e-31;
Matches 266; Conservative 160; Mismatches 408; Indels 198; Gaps 44;

QY 81 CFCNLLG-SFTVLGRGHSITFENIRTSNGAALS---NSAAGLFTIEGPKELSP-SNCN 135
DB 5 CFCNLLPFTFVLNAGLQPLEYIT-----LSPEYQAAPOVGFTHNQDLAIVGNHN 58
QY 136 SLIAVLPAATNKGSSQT-----PTTSTPNSGTIYSKTDLLLNKEKFP 180
DB 59 DFILDYKYRSNGGALTCKNLLISENIGNVFPEKNCVPSGGAIYAAQNCTISKQNYAF 118
QY 181 YSNLVSGD-----GGIDAASLTQVIGSKLCVQENTAQADGACQVTSFSAWAN 231
DB 119 TTNLVSDNPATAGSLIGALFAINCITNNLGGTFVDNLALNKGALTYETNLSKDN 178
QY 232 EAPITAFVANVA---GVRRGGGIAAVQDQGGQGVSTSTEDPVSFSRNTAVEFDGNVARVG 288
DB 179 KGPILIKQNRALNSDSLGGGIYS-----GNSLNIE-----GNSGAIQITSNNGSG 224
```

```
QY 289 GGIYSYGNVAFLNNGKTLFLNNVASPVYIAAQPTSGQASNTSNNYDGG-----AIFCN 344
DB 225 GGIFSTQTLTSSNKKLIEISE-----NSAFANNYGSNPNFGGGLTTTFC-- 270
QY 345 GAQAGSNNSGVSPDGEVGFSSNVAAGKGAIAKLSVANCPVOPFLNIANDGAI 404
DB 271 -----TILNREGVL-FNNNQSQSNGGAHAKSIIIKENGFPVFLNNTATRGAL 319
QY 405 YL-----GBSGLSLSADYDIIIFCNLKRXTAKENAAD--VNGVTVSSQAISGSGKITT 458
DB 320 LNLGSGNGSGFILLSADNGDIIIFNNN---TASKHALNPYRNAIHSTPMNLIQIG----- 371
QY 459 LRKAGHOILFNDPIEMANGNNQPAQSKLLKINDGEYGTGDIIVFANGSSTLYQNVT--- 515
DB 372 --ARPGYVLFDPIE-----HELPSFPILFNPETGHTGTGLFS--GEHVHONFTDEM 421
QY 516 -----IEQRIVLRKAKLSVNSLSQTGSLYMEAGS--TWDPVTQPQPQPPA 562
DB 422 NPFYSYLRNTSELROGVLAVEDGAGLACYKFFQGGTLLLGQGAIVTTAGTIPTPSTPTT 481
QY 563 ANQLITLNLHLSLSLLANNAVTPNPPPAQDGHPAVIGST----TAGSVTISGPIFP 618
DB 482 VGSTITLNLHIAIDLSILSFQA-----QAPKIWIYPTKGTSTYTEDSNPTTISGTLT 535
QY 619 EDL-DDTAYDYDMLGSNQKINVLKL-QLGTKPPANAPSDLLTGNEMPKYGGQSKLAW 676
DB 536 RNSNNEPDVSLDLSHLSLEKVPVLYVDVAAQKINSQDLSTLNSGEHYGYGGIWSITYW 595
QY 677 DNPT-----ANNPPTLKATWTKYGNPGBERVASLVPSNLWSILDIRSAHSAL 726
DB 596 VETTTITNPTSLLGANTKHKLLYANWSPILGYRPHRPERGFEITNALMQSAYTALAGLHSL 655
QY 727 QASVDGRSVCRLGWGVSNPFYHDDRALG-QGYVYISGGYSLGANSYFGSS-MFGLAFT 784
DB 656 SSWDEKHAASLOGIGL---LVHQDKNGFGRFSGHMTGYSATTATSSQSPNLSLGA 712
QY 785 EVFGRSKDYV--CRSNHACIGSVYLSLQALCGSYLFGDAFIRAS-----YGFQGNQHM 838
DB 713 QPFSKAKEHESQNSSTSSH---YFS---GMCIENTLFKEWIRLSVSLAYMFTSEHTH 763
QY 839 TSYTFAEE--SDVRDNDNCLAGEIGAGLPVITPS-----KLYLNELEPPFQAEFSYAD 890
DB 764 TMYQGLLEGNSQGSFHNHTLAG---ALSCVFLPQPHGESLQIY-----PFITALAIRGN 814
QY 891 HESFTEEGDQARAFKSGH--LLNLSPVGVKDRCSSTHPNKYSFM-----AAVYICDAYRT 944
DB 815 LAAPQESGDHAREP-SLHRPLTDVSLPVGI---RASWKHHRVPLVWLTEISYRSTLYRQ 870
QY 945 ISGTETTLTSHOETWTTDAFLARH--GVVVRGSM--YASLTSNIEVYGHGREYRDASR 1000
DB 871 DPELHSLKLLISQGTWTTQATPVYTNALGIKVKNTMQVPPKVTLSLDYSA---DISSSTL 926
QY 1001 GYGLSAGSRVRF 1012
DB 927 SHYLNVASRMRF 938

RESULT 37
H71460
probable outer membrane protein H - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: H71460
R:Stephens, R.S.; Kalmann, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Mitchell
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: H71460
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1016 <ARN>
A:Cross-references: UNIPROT:O84880; UNIPARC:UPI0000046AC7; GB:AE001360; GB:AE001273; NID
A:Experimental source: serotype D, strain UW-3/Cx
```



Db 413 PNPEHLGTLTLPSSAVIPDTSTGRDDPFISHFRNHGILYNGTLLALEDRAEKWKVYKFDQFGGT 472

Qy 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTVN---PPTN---PPA 594

Db 473 LRL--GSRVAVFTTDEQSSSVGSGVININNLAINPSILGNRVAPKWLIRPTGSSAPYS 530

Qy 595 QDSHPAVIGSTTAGSVTTISGPIFFEDLDATAYDRYDNLGSNQ---KINVLKQLQGTGKPPA 651

Db 531 EDNNPI-----INLSGPLSL--LDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHI 579

Qy 652 NAPSDLTLG-NEMPKYGYGSKWLAW-----DPNTANGPYTLKATTKTGYNPG 700

Db 580 NTDNFPYEGTLNTHQYGYGQVNSPYWIEITITSDTSSEDTVNTLHRLQYGDWTPGTGKYN 639

Qy 701 PERVASLVPNSLWGS-----LDIRSAHSAQASVDGRSYCRGLWVSGVSNFFYH--D 751

Db 640 PENKGDIALSAFWQSFHNLFATLYTQOQOQAPTASGEA-----TRLFVHQNSN 689

Qy 752 RDALGQYRYISGYSGLGANSYFGSS-MFGLAPTEVF-----GRSKDYVVCNRHHC-I 804

Db 690 NDA--KGFMEATGYSLGTTNTASNHSFGVNSQLFSNLYESHSDNSVASHSTTTVALQI 747

Qy 805 GSVELTQOALCGSYLFGDAPIRASYGFGNQHKMTSYTFAESDVRWNNCLAGETGAGL 864

Db 748 NNPWLQERFSTAS-----LAYSNSHHIKAS---GYSKIQTEGKCYSTTLGAAL 795

Qy 865 -----PIVITPSKLYLNELPFPVQAEFSYADHESFTTEGQARAFKSGH--LLN 911

Db 796 SCLSLQWRSRPLHFT-----PFTQAIARNSQNTAFQESGDKARKE-SVHKPLYN 844

Qy 912 LSPVGV-----KPDRCSSSTHPNKYSFMAAYICDAYRTISGTTETLLSHQETWTTDAFH 965

Db 845 LTVPLGIQSAWESKF-----RLPTYWNIELAYQVLYQQNPVNVSLSSGSSWLLSGIT 899

Qy 966 LARHGVVVGRGMYASLTNSIEVYGHGRYEDRASRGYGLSAGSRVRF 1012

Db 900 LARNAIAPKGRNQIFIPPKLSVFLDYQGSVSSSTTTTHYLHAGTTKPF 946

RESULT 39

DB1594

polymorphic membrane protein E/F family CP0283 [imported] - Chlamydomophila pneumoniae (se

C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004

C;Accession: D81594

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000

A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A;Reference number: AB1500; MUID:20150255; PMID:10684935

A;Accession: D81594

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-946 <REA>

A;Cross-references: UNIPROT:Q92880; UNIPARC:UPI00000131C86; GB:AE002161; NID

A;Experimental source: strain AR39, HL cells

A;Genetics:

A;Gene: CP0283

Query Match 11.6%; Score 609; DB 2; Length 946;

Best Local Similarity 24.7%; Pred. No. 5.2e-29;

Matches 264; Conservative 138; Mismatches 419; Indels 246; Gaps 43;

Qy 60 PSAGELTKNLNDNSIALPLS-CFGLNLSGFTVLGRGSLTFPENIRTSNGAALSNSAAD 118

Db 12 PFGALILKTKTILLNATPLSDYFDNQANQLTTL-----FPLIDLTNTWTPYSRA-- 62

Qy 119 GLFTIEGFKELSPNCNSLLAVLPAATNKGSGQPTTSTPSNGTIYSKTDLLILNNEKF 178

Db 63 ---TLFGVRD-----DTNQ-----DIVLDHONSI 83

Qy 179 SPYSNLVSGDGAIDAKSLTYQGISKLCVQFNTAQDGGACQVVTSPSAMANEAPAFV 238

## RESULT 40

C72075

polymorphic

outer membrane protein e/f family - Chlamydomophila pneumoniae (strain CWL029

C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: C72075

R;Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J

Nature Genet. 21, 385-389, 1999

A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A;Reference number: A72000; MUID:99206606; PMID:10192388

Db 84 ESWFENFSQDGGALSKCKSLAITNTQNLITLNSFAIKRAGAMYVNGNFDLSENHGSILPS 143

Qy 239 ANVAGVRGGGIAAVQDQGGQSSSTSTEDPVVVSFSRNTAVFPDGNVAVRGGIY----- 292

Db 144 GNLSFPNASPADTCTGGAVLCSKNVT-----ISKQGTAY-FINKAKSSGGAIQAAIIN 198

Qy 293 ---SYGNVAFPLNN-----GKTLFLN-----NVASPVYIAAKOPTSGQA----- 327

Db 199 IKDNTGCPCLFFNNAAGTAGGALFANACRIENNSQPIYPLNNQSGGLGGAIRVHQCILTK 258

Qy 328 -----SNTSNNTYGDGAIFCKGQAQAGSNNSGVSFQDGEVWFSSNVAAG 373

Db 259 NTGSVIFNNFMAMEADISANHSSGGAIYC-----ISCSIK-DNPGIAAFDNNNTAAR 308

Qy 374 KGAIAVAKKLSVANGCPVQFLNIANDGGAIVLBSGELSLSADYGDIIIFDGNLKRKTAKE 433

Db 309 DGAICTQSLTIQDSGPVFTNNQGTGGAIMLRQDGACTLFADQGDIIIFYNN--RHFKD 366

Qy 434 NAADVNGVTVS--SQAISMGSGGKIITLRKAGHQILFNDPIEMANGNNOQPAQSSKLLKIN 492

Db 367 TPS--NHVSVNCTRNVS-----TVGASQGHSAIFYDPIL-----QRYTIQNS-IOXFN 412

Qy 493 DGEYTGDIYFANG-----SSTLYQNVITIEQGRIVLREKAKLSVNSLSOTGGS 540

Db 413 PNPEHLGTLTLPSSAVIPDTSTGRDDPFISHFRNHGILYNGTLLALEDRAEKWKVYKFDQFGGT 472

Qy 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTVN---PPTN---PPA 594

Db 473 LRL--GSRVAVFTTDEQSSSVGSGVININNLAINPSILGNRVAPKWLIRPTGSSAPYS 530

Qy 595 QDSHPAVIGSTTAGSVTTISGPIFFEDLDATAYDRYDNLGSNQ---KINVLKQLQGTGKPPA 651

Db 531 EDNNPI-----INLSGPLSL--LDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHI 579

Qy 652 NAPSDLTLG-NEMPKYGYGSKWLAW-----DPNTANGPYTLKATTKTGYNPG 700

Db 580 NTDNFPYEGTLNTHQYGYGQVNSPYWIEITITSDTSSEDTVNTLHRLQYGDWTPGTGKYN 639

Qy 701 PERVASLVPNSLWGS-----LDIRSAHSAQASVDGRSYCRGLWVSGVSNFFYH--D 751

Db 640 PENKGDIALSAFWQSFHNLFATLYTQOQOQAPTASGEA-----TRLFVHQNSN 689

Qy 752 RDALGQYRYISGYSGLGANSYFGSS-MFGLAPTEVF-----GRSKDYVVCNRHHC-I 804

Db 690 NDA--KGFMEATGYSLGTTNTASNHSFGVNSQLFSNLYESHSDNSVASHSTTTVALQI 747

Qy 805 GSVELTQOALCGSYLFGDAPIRASYGFGNQHKMTSYTFAESDVRWNNCLAGETGAGL 864

Db 748 NNPWLQERFSTAS-----LAYSNSHHIKAS---GYSKIQTEGKCYSTTLGAAL 795

Qy 865 -----PIVITPSKLYLNELPFPVQAEFSYADHESFTTEGQARAFKSGH--LLN 911

Db 796 SCLSLQWRSRPLHFT-----PFTQAIARNSQNTAFQESGDKARKE-SVHKPLYN 844

Qy 912 LSPVGV-----KPDRCSSSTHPNKYSFMAAYICDAYRTISGTTETLLSHQETWTTDAFH 965

Db 845 LTVPLGIQSAWESKF-----RLPTYWNIELAYQVLYQQNPVNVSLSSGSSWLLSGIT 899

Qy 966 LARHGVVVGRGMYASLTNSIEVYGHGRYEDRASRGYGLSAGSRVRF 1012

Db 900 LARNAIAPKGRNQIFIPPKLSVFLDYQGSVSSSTTTTHYLHAGTTKPF 946

A:Accession: C72075  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-946 <ARN>  
A:Cross-references: UNIPROT:Q9Z880; UNIPARC:UPI0000165625; GB:AE001631; GB:AE001363; NID  
A:Experimental source: strain CWL029  
C:Genetics:  
A:Gene: pmp\_18

Query Match 11.5%; Score 607; DB 2; Length 946;  
Best Local Similarity 24.8%; Pred. No. 6.9e-29;  
Matches 265; Conservative 139; Mismatches 417; Indels 246; Gaps 44;

QY 60 PSAGELTLKLNLDNSTAALPLS-CFQNLGLSFVILGRGHSILTPENTRTSTNGAALSAAAD 118  
DB 12 FVFGALILGKTTILLNATPLSDYPONQANQLTTL-----PPLIDTLTNMTPYSHRA-- 62

QY 119 GLFTIEGPKELSFSCNSLLAVLPAATTNKGSTPTTTSTPSNGTIYSKTDLLLNNEKF 178  
DB 63 ---TLFGYRD-----DTNQ-----DIVLDHQNSI 83

QY 179 SFYSNLVSGDGAIDAKSLTVQIGSKLCVFQENTAQADGGACQVVTSPSAMANEAPIAPV 238  
DB 84 ESWFENFSQDGGALCKSLAITNTKQILFLNSPAIKRAGAMYVNGNFDLSENHGSIIIFS 143

QY 239 ANVAGVGGGTAIVODGQGVSSSTSTEDPVVSPSNTAVFEDGNVAVGGGIY----- 292  
DB 144 GNLSPFNASNPADTCTGGAVLCCKNVT-----ISKQGTAY-FINNKAUSSGGAIOAAIIN 198

QY 293 ---SYGNVAFLLN-----GKTLFLN-----NVASPVYIAAKQPTSGQA----- 327  
DB 199 LKNTGCPCLFFNNAAGGTAGGALFANACRIENNSQPIYFLNQSGLGGAIRVHQECILTK 258

QY 328 -----SNTSNNGYDGGAIKFCNGAQAAGSNNSVSVDGEGVVFPPSSNVAAG 373  
DB 259 NTGSVIFNNNPAMEADIGANSHSGGAIYC-----ISCSIK-DNPGIAAFDNNNTAAR 308

QY 374 KGGAIYAKKLVANGCPVQPLNIANDGGAIVLGESGELSADYGDIIIPGNLKRITAKE 433  
DB 309 DGGAICTQSLTTIQDSGPVYFTNNQGTWGGAIMLRDQGACTLFDQGDIIIFYNN--RHPKD 366

QY 434 NAADVNGVTVS--SOAISHMSGGKITTAKAGHQILFNDPIEMANGNNOQAQSSKLLKIN 492  
DB 367 TFS--NHVSVNCTRNVS-----TVGASQGSATFYDPIL-----QRYTIONS-IQKFN 412

QY 493 DDEGYTGDIVFAN-----GSST-----LYQNVITIEQRIVIREKAKLSVNSLSQTGGS 540  
DB 413 PNPEHLGTILFSSTVIPTDTSRDDFISHFRNHIGLYNGTLALEDRAEMKVYKPDQFGGT 472

QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNVNTN---PPTN---PPA 594  
DB 473 LRL--GSRVFTTDEQSSSSVSGSVINNLAINLPSILGNRVAPKLWIRPTGSSAFYS 530

QY 595 QDSHPAVIGSTTAGSVTISGPFFEDDDTAYDRYDMLGSNQ---KINVKLQLGTKPPA 651  
DB 531 EDNPI-----INLSGPLSL--LDENLDPYDTADLAQPIAEVPELLYLLDVTAKHI 579

QY 652 NAPSDLTLG-NEMPKYIGQGSWKLAW-----DPNTANNPGPYTLKATWTKTGYNPG 700  
DB 580 NTDNFPYEGLNTHQHYGQGVWSPYWIETITTSSTSSDVTNLTLRQLYGDWTPGTGYKN 639

QY 701 PERVASLVPNSLWGI-----LDIRSAHSAIQASVDSYCRGLWVSGVSNFFYH----D 751  
DB 640 PENKGDIALSAPWQSFHNLPAFLRYCTQOQGIAPTASGEA-----TLFVHQNSN 689

QY 752 RDALQGVYITSGGYSLGANSYFGSS-MFGLAFTEVF-----GRSKDYVVCRSNHHAC-I 804  
DB 690 NDA--KGFHMEATGSLGTTNTASNHSFGVNFSQLFNLYESHSDNSVASHTTVALQI 747

QY 805 GSVYLSQQAALCGSLFGDAFIRASVFGNQHMKTSYTFABESDVWRDNNCLAGEIGAGL 864  
DB 748 NNPMLOERFSTAS-----LAYSYSNHHIKAS---GYSGKIQTGCKVSTTLGAAL 795

QY 865 -----PIVITPSKLYNLNRPVQAEFSYADHESFTEBGDQARAPKSGH--LLN 911  
DB 796 SCSLSLQWRSRPLHPT-----PFIQAIAVRSNQTAFQESGDKARKF-SVHKPLYN 844

QY 912 LSVPVGV-----XFDRCSSSTHPNKNYSFMAAYICDAVRTISGTETTTLLSHQBTWTTDAPH 965  
DB 845 LTVPLGIQSAMESKF-----RLPTYWNIELAYQPVLYQONPEINVSLESSGSSWLLSGTT 899

QY 966 LARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF 1012  
DB 900 LARNAIAFKGRNQIFIPFKLSVFLDYQGSVSSSTTTHYLHAGTTPKF 946

Search completed: May 13, 2006, 12:20:38  
Job time : 62 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:15:19 ; Search time 243 Seconds  
(without alignments)  
2938.250 Million cell updates/sec

Title: US-10-701-844-2

Perfect score: 5267

Sequence: 1 MQTSHKFLSMILAYSCS.....YETDASRGYGLSAGSRVRP 1012

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5131.5	97.4	1013	1	PMPG_CHLTR
2	3845	73.0	987	1	Q9P145 chlamydia t
3	1860.5	35.3	1011	2	Q823X5 CHLCV
4	1776	33.7	1024	2	Q5L6J7 CHLAB
5	1365.5	25.9	973	1	PMP13 CHLPN
6	1158	22.0	942	2	Q823X1 CHLCV
7	1122	21.3	926	2	P71135 CHLAB
8	1122	21.3	926	2	Q5L6J3 CHLAB
9	1101	20.9	928	1	PMP10 CHLPN
10	1099	20.9	926	2	Q823W9 CHLCV
11	1090	20.7	928	1	PMP11 CHLPN
12	1079	20.5	930	1	PMP8 CHLPN
13	1077	20.4	1378	2	Q5L6J2 CHLAB
14	1040	19.7	936	1	PMP7 CHLPN
15	1021	19.4	928	1	PMP9 CHLPN
16	1012	19.2	1276	1	PMP6 CHLPN
17	1008	19.1	866	2	Q823X0 CHLCV
18	978.5	18.6	772	2	Q9RB71 CHLPN
19	971.5	18.4	841	2	Q822Q5 CHLCV
20	967	18.4	922	1	PMP1 CHLPN
21	950	18.0	847	2	P71132 CHLAB
22	949.5	18.0	839	2	P77792 CHLAB
23	939.5	17.8	841	1	PMP2 CHLPN
24	897	17.0	846	2	P71133 CHLAB
25	803	15.2	843	2	Q823X2 CHLCV
26	789	15.0	867	1	PMP1 CHLMU
27	787.5	15.0	602	2	Q8VU49 CHLPS
28	785.5	14.9	843	2	Q823X3 CHLCV
29	785	14.9	849	2	Q5L6J5 CHLAB
30	784.5	14.9	878	2	Q4VWS6 CHLTR
31	781.5	14.8	878	2	Q4VWR2_CHLTR

32	779.5	14.8	878	2	Q4VWS3_CHLTR	Q4VWS3 chlamydia t
33	779.5	14.8	878	2	Q4VWS8_CHLTR	Q4VWS8 chlamydia t
34	777.5	14.8	878	2	Q4VWR8_CHLTR	Q4VWR8 chlamydia t
35	773.5	14.7	840	2	Q5L6J6_CHLAB	Q5L6J6 chlamydia t
36	771.5	14.6	846	2	Q84FU4_CHLTR	Q84FU4 chlamydia t
37	771.5	14.6	878	1	PMP1_CHLTR	Q84882 chlamydia t
38	771.5	14.6	878	2	Q4VWS2_CHLTR	Q4VWS2 chlamydia t
39	770.5	14.6	878	2	Q4VWR1_CHLTR	Q4VWR1 chlamydia t
40	769.5	14.6	846	2	Q84FU2_CHLTR	Q84FU2 chlamydia t
41	769.5	14.6	878	2	Q4VWR6_CHLTR	Q4VWR6 chlamydia t
42	766.5	14.6	846	2	Q84FU1_CHLTR	Q84FU1 chlamydia t
43	766.5	14.6	846	2	Q84FU3_CHLTR	Q84FU3 chlamydia t
44	764.5	14.5	846	2	Q84FU0_CHLTR	Q84FU0 chlamydia t
45	764.5	14.5	846	2	Q84FU5_CHLTR	Q84FU5 chlamydia t

## ALIGNMENTS

### RESULT 1

PMPG\_CHLTR  
ID\_PMPG\_CHLTR STANDARD; PRT; 1013 AA.  
AC O84879;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G).  
DE GN Name=pmpG; OrderedLocusNames=CT871;  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=D/UM-3/Cx;  
RX MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;  
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;  
RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";  
RT Science 282:754-759(1998).  
RL -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).  
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.  
CC -!- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AR001360; AAC68469.1; -; Genomic\_DNA.  
DR PIR; G71460; G71460.  
DR PHCI-2DPAGE; O84879; -  
DR InterPro; IPR005546; Auto\_transptbeta.  
DR InterPro; IPR006315; Autotransporter.  
DR InterPro; IPR011427; ChlamPMP\_M.  
DR InterPro; IPR003368; Chlamydia\_PMP.  
DR Pfam; PF03797; Autotransporter; 1.  
DR Pfam; PF02415; Chlam\_PMP; 9.  
DR Pfam; PF07548; ChlamPMP\_M; 1.  
DR TIGRFAMs; TIGR01414; autotrans\_bar1; 1.  
DR TIGRFAMs; TIGR01376; POMP\_repeat; 5.  
DR Complete proteome; Membrane; Multigene family; Outer membrane; Signal.  
FT SIGNAL 1 27 Potential.  
FT CHAIN 28 1013 Probable outer membrane protein pmpG.  
SQ SEQUENCE 1013 AA; 107367 MW; F0927743C0A651DD CRC64;  
Query Match 97.4%; Score 5131.5; DB 1; Length 1013;  
Best Local Similarity 97.0%; Pred. NO. 3.1e-295;  
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY	1	MQTSFHKFLSHILAYSCSLNGGYYAAIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP	60
DB	1	MQTSFHKFLSHILAYSCSLNGGYYAAIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP	60
QY	61	SAGELTLKNLONSIAALPLSCFNLGSLTFVLGRHSITFENIRTSNGAALSNSAADGL	120
DB	61	SAGELTLKNLONSIAALPLSCFNLGSLTFVLGRHSITFENIRTSNGAALSNSAADGL	120
QY	121	FTIEGFKELSPNCNLSLAVLPAAATNKGSTPTTSTPSNGTIYSKTDLLLNNKFSF	180
DB	121	FTIEGFKELSPNCNLSLAVLPAAATNKGSTPTTSTPSNGTIYSKTDLLLNNKFSF	180
QY	181	YENLVSGDGAIDAKSLTVQISKLCPQENTQAADGGACQVTVTSFSAANEAPIAFVA	240
DB	181	YENLVSGDGAIDAKSLTVQISKLCPQENTQAADGGACQVTVTSFSAANEAPIAFVA	240
QY	241	VAGVRGGGIAAVQDQGGVSSSTSTEDPVWFSRNTAVFPGNVARVGGGIYSYGNVAF	300
DB	241	VAGVRGGGIAAVQDQGGVSSSTSTEDPVWFSRNTAVFPGNVARVGGGIYSYGNVAF	300
QY	301	NGKTLFLNNVSPYIAAKQPTSGASNTSNNGYDGGAIKFCNGAQ-AGSNNGSVSFD	359
DB	301	NGKTLFLNNVSPYIAAKQPTSGASNTSNNGYDGGAIKFCNGAQAGSNNGSVSFD	360
QY	360	GGVVFSSNVAAGKGAIYAKLSVANGCPVQFLRNTANDGGAITYGSELSLSADYG	419
DB	361	GGVVFSSNVAAGKGAIYAKLSVANGCPVQFLRNTANDGGAITYGSELSLSADYG	420
QY	420	DIIPGNLKRTAKENAAVNGVTVSSQAIKSGSGKITTLAKAGHOLLFNDPIEMANGN	479
DB	421	DIIPGNLKRTAKENAAVNGVTVSSQAIKSGSGKITTLAKAGHOLLFNDPIEMANGN	480
QY	480	NQAQSSKLLKINDGEGTGDIVFANGSSTLYQNTIEQGRIVLREKAKLSVNSLSQTGG	539
DB	481	NQAQSSKLLKINDGEGTGDIVFANGSSTLYQNTIEQGRIVLREKAKLSVNSLSQTGG	540
QY	540	SLYMEAGSTWDPVTPPQQPPAANQLITLNLHLSSLSLANNVTPNPQADSHP	599
DB	541	SLYMEAGSTWDPVTPPQQPPAANQLITLNLHLSSLSLANNVTPNPQADSHP	600
QY	600	AVIGSTTAGSVTISGPIFEDLDDTAYDRYDNLGSKNIVLKLQGTGKPPANAPSDTL	659
DB	601	AVIGSTTAGSVTISGPIFEDLDDTAYDRYDNLGSKNIVLKLQGTGKPPANAPSDTL	660
QY	660	QNEPKYQGSWKLAMPNTANNPGLTKATWTKTGNPQPERVASLVPNSLWGSILDI	719
DB	661	QNEPKYQGSWKLAMPNTANNPGLTKATWTKTGNPQPERVASLVPNSLWGSILDI	720
QY	720	RSASIAIOASVDRGSRGSLVGVSNFPHYDRDALQGYRYISGGYISLGANSYFGSSMP	779
DB	721	RSASIAIOASVDRGSRGSLVGVSNFPHYDRDALQGYRYISGGYISLGANSYFGSSMP	780
QY	780	GLAPTEVFGSKDYVVCSSNHACIGSVYLSLQALCGSLYFGDAPFASVGRGNHMT	839
DB	781	GLAPTEVFGSKDYVVCSSNHACIGSVYLSLQALCGSLYFGDAPFASVGRGNHMT	840
QY	840	SYTFAESDVRWNNCLAGEICAGLPVITPSKLYLNEIRFPVQAEFSDYADHESFTEGD	899
DB	841	SYTFAESDVRWNNCLAGEICAGLPVITPSKLYLNEIRFPVQAEFSDYADHESFTEGD	900
QY	900	QARAFKSHLNLVSVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	959
DB	901	QARAFKSHLNLVSVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	960
QY	960	TTDAPHLARHGIVVGRGSMYASITSNIEVYHGRVEYRDASRGYGLSAGSVRP	1012
DB	961	TTDAPHLARHGIVVGRGSMYASITSNIEVYHGRVEYRDSRGYGLSAGSVRP	1013
RESULT 2			
PMPG_CHLWU			
ID	PMPG_CHLWU	STANDARD;	PRT; 987 AA.

AC	Q9PL45;	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G).		
GN	Name=pmpG; OrderedLocusNames=TC0263;		
OS	Chlamydia muridarum.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=83560;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	STRAIN=MoPn / N199;		
RX	MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;		
RA	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,		
RA	White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,		
RA	Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,		
RA	Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,		
RA	McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.; Chlamydia		
RT	pneumoniae AR39.;		
RL	Nucleic Acids Res. 28:1397-1406(2000).		
CC	-1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)		
CC	(Potential).		
CC	-1- SIMILARITY: Belongs to the PMP outer membrane protein family.		
CC	-1- SIMILARITY: Contains 1 autotransporter (TC 1.9.12) domain.		
CC	-----		
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use as long as its content is in no way modified and this statement is not		
CC	removed.		
CC	-----		
DR	EMBL; AB002293; AAF39132.1; -; Genomic_DNA.		
DR	PIR; H81722; H81722.		
DR	TIGR; TC0263;		
DR	InterPro; IPR005546; Auto_transp_beta.		
DR	InterPro; IPR006315; Autotransporter.		
DR	InterPro; IPR011427; ChlamPMP_M.		
DR	InterPro; IPR003368; Chlamydia_PMP.		
DR	Pfam; PF02415; Autotransporter; 1.		
DR	Pfam; PF07548; ChlamPMP_M; 1.		
DR	TIGRFAMs; TIGR01414; autotrans_bar1; 1.		
DR	TIGRFAMs; TIGR01376; POMP_repeat; 5.		
DR	Complete proteome; Membrane; Multigene family; Outer membrane; Signal.		
KW	Potential.		
FT	SIGNAL 1 25		
FT	CHAIN 26 987 Probable outer membrane protein pmpG.		
FT	SEQUENCE 987 AA; 104867 MW; 32079BD6BEB2DA42 CRC64;		
QY	Query Match	73.0%; Score 3845; DB 1; Length 987;	
DB	Best Local Similarity	72.6%; Pred. No. 5.1e-219;	
DB	Matches	739; Conservative 105; Mismatches 136; Indels 38; Gaps 13;	
QY	1	MQTSFHKFLSHILAYSCSLNGGYYAAIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP	60
DB	2	MQTSFHKFLSHILAYSCSLNGGYYAAIMVPOGIYDGETLTVSPFYTVIGDPSGTTVP	58
QY	61	SAGELTLKNLONSIAALPLSCFNLGSLTFVLGRHSITFENIRTSNGAALSNSAADGL	120
DB	59	SSGSELEKLNLSIATPLSCFNLGSLTFVLGRHSITFENIRTSNGAALSNSAADGL	118
QY	121	FTIEGFKELSPNCNLSLAVLPAAATNKGSTPTTSTPSNGTIYSKTDLLLNNKFSF	179
DB	119	FVIRAFDELILNLSVSVVP-----QTGGTTTSPVSNGLIYSRTDLVLRDIKVS	170
QY	180	FYSNLVSGDGAIDAKSLTVQISKLCPQENTQAADGGACQVTVTSFSAANEAPIAFVA	239
DB	171	FYSNLVSGDGAIDAKSLTVQISKLCPQENTQAADGGACQVTVTSFSAANEAPIAFVA	230
QY	240	NVAGVRGGGIAAVQDQGGVSSSTSTEDPVWFSRNTAVFPGNVARVGGGIYSYGNVAF	299
DB	231	NVAGVRGGGIAAVQDQGGVSSSTSTEDPVWFSRNTAVFPGNVARVGGGIYSYGNVAF	286

```
QY 300 LNNKXTLFLNNVASPVYI--AAKQPTSGQASNTSNNGYGGGAIKCKNGAAGSNNNGSVS 357
DB 287 LGNAKTVFLSNVASIYYDPA--AGGQPPADKDNNGYGGGAIKCKN-----DTNIGSVS 338
QY 358 PDGEGWFFSSNVAAGKGAIIYAKKLSVANCGPVQFLRNANDGAIYLGESGELSAD 417
DB 339 PDGEGWFFSKNIAAGKGAIIYAKKLTSDCGPVQFLGNVANDGAIYLDQGELSAD 398
QY 418 YGDIIFDGNLKTAKENADVNGTVSSQAIISNGSGGKITTLRAKAGHOILFNDPIEMAN 477
DB 399 RGDIIIFDGNLKRMAATGAATVHDVNVASNAISWATGGQITTLRAKEGRILFNDPIEMAN 458
QY 478 GNNQPAQSCKLKLKINDGEGYCDIIVPANGSSTLYQNVTIEQGRIVILREKAKLSVNSLSQT 537
DB 459 G--QPV--IQTLTVNEGEGYTDIIVFAKGDNDVLYSIELSQRIILREQTKLVNSLQT 514
QY 538 GGSGLYMEAGSTWDFVTPQPPQPPAANQLITLSNHLHLSLSLLANNAVTPNTPPAQDS 597
DB 515 GGSVMEGGSTLDFAV----TTPPAANS-ALTNVHFSLASLLKKNVTPNTPNPVQVS 569
QY 598 HPVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKNOKINVLKQLQGTKPPANAPSDL 657
DB 570 SPVIGNTAAGTVTISGPIFFEDLDDTAYDNNQWLGADQTDVQLHLGANPPANAPTDL 629
QY 658 TLGNEMPKYGYQGSWKLAWDPNTAN---NGPYTLKATWTKGVNPGPERVASI.VPNSLWG 714
DB 630 TLGNESKKGYQGSWTLQWEPDPPQPNNSYMLKASWTKTGVNPGPERVASI.VPNSLWG 689
QY 715 SILDTRSAHSAQASVDGRSICRGLWVGVSFPFYHDDRALGQVRYISGYSIGANSYF 774
DB 690 SILDVRSASHAQASIDGRAYCRGWIISGIFSYHDDQALGQYRHSIGYSIGANSYF 749
QY 775 GSSMFLGAPTEVFGRSKDYVVCRSNHACIGSVYLSSTQALCGSYLFGDAPTRASYGFGN 834
DB 750 GSSMFLGAPTEVFGRSKDYVVCRSNDHTCVGSVYLSSTQALCGSYLFGDAPTRASYGFGN 809
QY 835 QHMKTSYTPAESDVTWDDNCLAGEICAGLPIVITPSKLYLNLRLPVOAEPSYADHESP 894
DB 810 QHMKTSYTPAESNVRWDNNVGVGVGAGLPIMLAASKLYLNLRLPVOAEPSYADHESP 869
QY 895 TREGDQARAFKSHLANLSVPVGVKFDRCSSHTPNKYSFMAAYICDAYRTISGTETTLAS 954
DB 870 TREGDQARAFKSHLANLSIPVGVKFDRCSSKHPNKYSFMAAYICDAYRSISGTETTLAS 929
QY 955 HQTWTWTDAPHLARHGVMVGRGSMYASLTNIEYVGHRYEYRDASRGYGLSAGSRVRP 1012
DB 930 HKTWTWTDAPHLARHGVMVGRGSMYASLTNIEYVGHRYEYRDASRGYGLSAGSRVRP 987

RESULT 3
ID Q823X5 CHLCV PRELIMINARY; PRT; 1011 AA.
AC Q823X5
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family
DE protein/autotransporter.
GN OrderedLocusNames=CCA00278;
OS Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapfel E.K., Khouri H.M., Federova N.B.,
RA Carthy H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
RA Bavolli P.M., Fraser C.M.;
RT "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):
```

```
RT examining the role of niche-specific genes in the evolution of the
RL Chlamydiaceae.";
RE Nucleic Acids Res. 31:2134-2147 (2003).
DR EMBL; AB016995; AAP05029.1; -, Genomic_DNA.
DR TIGR; CCA00278; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005546; Auto transportbeta.
DR InterPro; IPR011427; Chlamydia_PMP.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; Chlamydia_PMP; 1.
DR Pfam; PF02415; Chlamydia_PMP; 7.
DR TIGRPFAMs; TIGR01414; autotrans baz1; 1.
DR TIGRPFAMs; TIGR01376; POMP_repeat; 6.
KW Complete proteome.
SQ SEQUENCE 1011 AA; 107952 MW; EBD9A9FALC46EDBE6 CRC64;

Query Match 35.3%; Score 1860.5; DB 2; Length 1011;
Best Local Similarity 39.6%; Pred. No. 2.1e-101;
Matches 426; Conservative 168; Mismatches 353; Indels 129; Gaps 27;

QY 1 MQTSFKFELS--MLAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPRYTVIGDPSGTT 58
DB 1 MKASLRKFLISTLTLPYSF-----QAFSLVVVPNGTYDG-NLRETPTTITSNEGTT 54
QY 59 VFSAGELTKNLDNSIAALPLSCFNLGSLFTVLRGHSHTFENIRTSNGAALSN---- 114
DB 55 AILSGNVLNLDNSNVAITPSCFPNSAGSMITVGRNHLITNLTNLSANGAALSIPIT 114
QY 115 SAADGLFTIEGPKELSPNCNSLLAVLPAATNKGSOPTTTTSTPSNGTIYKTDLLLN 174
DB 115 TPESPPYTIKGVNTLSPSNC---LALMARTTTAPNTTTPV---NPNGGAFYSKAPVLEN 168
QY 175 NKEKFPYSLVSGDGAIDAKSLTVQGTISKLQFQENTAAQDGGACQVTVTSAMANEAP 234
DB 169 IQNVLFKNRAADSGGLWVETAGISNKKSMQFLSNVG-ANGGA--INAKSLDVTQCP 225
QY 235 -TAFVANVAGVRGGGIAAVODGQGVSSSTSTEDPVVSPSRNTAVEFDGNVARVGGIYS 293
DB 226 SILFNSNSAEKUGGAIQAVD-----PATNQVNTAVRFSNGSVQFDANNAKSGAIYS 279
QY 294 YGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYGGGAIKCKN----- 342
DB 280 KGNVDFSNNAQLLIQNNSASP-----EVANTNEVLGGGAIKCKN----- 328
QY 343 ----KNGAAGSNNNGSVFDEGVVFPSSNVAAGKGAIIYAKKLSVANCGPVQFLRNIA 398
DB 329 PPPTTNPVPSG-----LTITNQKQILFANNFAATAGGAIYGEKVSITSGKTMFTNNA 382
QY 399 NDGGAIYLGESGELSADYGDIIIPDGNLKTAKENADVNGTVSSQAIISNGSGGKITT 458
DB 383 KDGGAIYTPENGELTSLASDYGDIMFYENLKK-----DDATVTRNAVTLAKGATIKL 433
QY 459 LRAKAGHOILFNDPIEMANGNNGNPQAOSKLLKINDGE-----GYTGDIVF----- 503
DB 434 LAASGDHKLCFYDPIVITLTPETAPT-NDKTLTINQDKTSSTPTTNYIGTLFPFGAVDSQ 492
QY 504 ----ANGSSTLYQNVTIEQGRIVILREKAKLSVNSLSQ-TGGSLYMEAGSTWDFTPQPP 557
DB 493 SASSTTANFESTIYQKIVILGGKLVLDKASLSVASFTQETSDTSLMDNGTT--LAITEHS 550
QY 558 QQPPAA-----NQLTLNHLHLSLSLLANNAVTPNTPPAQDSHPAV 601
DB 551 HOTPAAAGGGGGGGTPTQEBANTDGVISLNLHVNISLSLTQGGGAKLET----- 599
QY 602 IGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKN-KINVLKQLG--TKPPANAPSDLT 658
DB 600 --KNTDGTITLTHGVSLDDVSGTAYENHDLFNKQVTITNLSLSTAGDSKTTING-LDLT 656
QY 659 L-GNEMPKYGYQGSWKLAWDPNTANNGPYTLKATWTKGVNPGPERVASI.VPNSLWGSIL 717
DB 657 LRGDABEPQYGYQGSWQLAWENGADANKQKILKATWTKGTPTNPERQASLVPSNLWGAFI 716
```





RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";   
 RL Nat. Genet. 21:385-389(1999).   
 RN (2)   
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].   
 RC STRAIN=AR39;   
 RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;   
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,   
 RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,   
 RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,   
 RA Dodson R.J., Gwin M.L., Nelson W.C., Deboy R.T., Kolonay J.F.,   
 RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;   
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia   
 RT pneumoniae AR39.";   
 RL Nucleic Acids Res. 28:1397-1406(2000).   
 RN (3)   
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].   
 RC STRAIN=J138;   
 RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;   
 RA Shirai M., Hirakawa H., Kinoto M., Tabuchi M., Kishi F., Ouchi K.,   
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;   
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138   
 RT from Japan and CWL029 from USA.";   
 RL Nucleic Acids Res. 28:2311-2314(2000).   
 RN (4)   
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].   
 RC STRAIN=TW-183;   
 RX MEDLINE=20007584; PubMed=10539856;   
 RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,   
 RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;   
 RT "Molecular biology of Chlamydia pneumoniae surface proteins and their   
 RT role in immunopathogenicity.";   
 RL Am. Heart J. 138:S491-S495(1999).   
 CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)   
 CC (Potential).   
 CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.   
 CC -1- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.   
 CC -----   
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration   
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -   
 CC the European Bioinformatics Institute. There are no restrictions on its   
 CC use as long as its content is in no way modified and this statement is not   
 CC removed.   
 CC -----   
 CC EMBL; AE001629; AAD18595.1; -; Genomic DNA.   
 CC EMBL; AE002191; AAP38156.1; ALT INIT; Genomic DNA.   
 CC EMBL; BA000008; BAA98460.1; -; Genomic DNA.   
 CC EMBL; AE017158; AAP98401.1; -; Genomic DNA.   
 CC EMBL; AJ133034; CAB37074.1; -; Genomic DNA.   
 CC PIR; B86547; B86547.   
 CC PIR; F72076; F72076.   
 CC PHCI-2DPAGE; Q92896; -.   
 CC TIGR; CP0299; -.   
 CC InterPro; IPR005546; Auto\_transp\_beta.   
 CC InterPro; IPR011427; ChlamPMP\_M.   
 CC InterPro; IPR003368; Chlamydia\_PMP.   
 CC Pfam; PF02415; Autotransporter; 1.   
 CC Pfam; PF07548; Chlam\_PMP; 8.   
 CC TIGRFAMs; TIGR01376; POMP\_repeat; 6.   
 CC Complete proteome; Membrane; Multigene family; Outer membrane; Signal.   
 FT SIGNAL 1 24 Potential.   
 FT CHAIN 25 973 Probable outer membrane protein pmp13.   
 FT CONFLICT 258 258 N -> Y (in Ref. 5).   
 FT SEQUENCE 973 AA; 102762 MW; E02A69F611DEBFE2 CRC64;   
 SQ

Query Match 25.9%; Score 1365.5; DB 1; Length 973;   
 Best Local Similarity 33.5%; Pred. No. 4.2e-72;   
 Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;   
 QY 1 MQTSHFPLSMILAYSCCSLNGGYYAAEIMVPGIYDGETLTVSPFVTVIGDPSGTTVP 60   
 DB 1 MKTSIRKFLSTTLAPCFASFT--AFTVEVIMPENFDGSGGKI-PPITLSDRGTLICI 56   
 QY 61 SAGSLTLKNDLNSIAALPLSCFGLLGSFTVLGRHSLTFENIRTSNTGAALSAAADG- 119   
 DB 57 PSGLDYLANLDNALSRSSSCFCFNAGALQILKGKGVPSFLNIRSSDGAALSSVITQNP 116   
 QY 120 ---LFTIEGFKELSFNSCNLLAVLPAATNKGSGTPTTSTPNSNGTIYSKTDLLLNNE 176   
 DB 117 ELCPLSPFSQMIFFDNCSLT-----SDTSASNVIPHASAIYATPMLFTND 165   
 QY 177 KPSFYNLVSGDGAIDAKSLTVQISKL CVFQENTAAQDGGACQVVTSPSAMANEAPIA 236   
 DB 166 SILFQYNSAGFGAAIRGTSITIENTKSLLFNGNGSISNGGALTGSAAIINLNSAPVI 225   
 QY 237 FVANVAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFEDGNVARVGGGIYSYGN 296   
 DB 226 FSTNATGIYGAIVLT-----GGSLTIS-----GNLSGVLVFNSSRSRGAIYANGN 272   
 QY 297 VAFLNNGKTLFLNNVASPV-YIAAKQPTSGQASNTNNYGDGGAIFCKNGQAQAGSNNGS 355   
 DB 273 VTFNNSDLTFQNTASTSPQNSLPAPTPTTTPAVTP-LLGYGGAIFCTPPATPTTGV-S 330   
 QY 356 VSPDGEVVFSSNVAAGKGAIAKLVANCGVPQVPLNIANDGGAIYLGESGELSLS 415   
 DB 331 LTISGNSVTFLENIASQGGALYKKISIDSNKSTIFLNTAGKGAIAIPESGELSLS 390   
 QY 416 ADYGDIIFDGNLKR TAKENAADVNGVTVSSQASISMGSGKITTIRAKAGHQLFNDPIEM 475   
 DB 391 ANQGDILFNKMLSTSG-----TFTRNSIHFGKDAKATLCATQGYTLYFYDPI--- 439   
 QY 476 ANGNQPAQSSKLLKIN-----DGEYTGDIIVF-----ANGSSTLYQNVTIE 517   
 DB 440 TSDDLASAASAAATVVVNFKASADG-AYSGTIVFSGETLTATTAATPATANATLNQKLE 498   
 QY 518 QGRIVLRKAKLVNSLSQTGGS-LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSL 576   
 DB 499 GCTALRNGATLVNHFVTQDEKSVVINDAGTT--LATNGANNTDGA---ITLNLKVLNL 553   
 QY 577 SLLANNAVTPPTPPAQDHPAIVIG-STTAGSVTISGPIPFEDLDLDTAVDRYDNLGSN 635   
 DB 554 DSLDGTKA-----AVNVQSTNGALTIISGLVKNQSCDCCDHGMFKND 598   
 QY 636 -QKINVLKLQGTQPPANAPSDL-TLGNEMPKYGYGGSWKLAWDPNTANGPYTLKATWT 693   
 DB 599 LQQVILELKATSNVTVTTFDPSLTNGYQSPYQGTWFTTIDTTT-----HTVTGNWK 653   
 QY 694 KTYNGPGERVASLPVNSLWMSILDIRSAHSIAQASVDGRSY-CRGLWVSGVSNFYHNR 752   
 DB 654 KTGYPHPERLAPLIPNSLWANVILDLRAVSQA--SAADGEDVPGKQLSITGITNPFANH 711   
 QY 753 DALGQYRIVSGYSIGANSY---FGSSMFGLAFTVEGRSKDYVVCNRRHACIGSVYL 809   
 DB 712 TGDARSYRHMGGGYLI--NTYTRITPDAAISLGFQQLFTKSKDYLHGHGHSNVYFATVVS 769   
 QY 810 STQOALCG-SYLFQDAPFIRASVGFQGNHMKTSYTFABESDVRWNNCLAGEIGAGLPYVI 868   
 DB 770 NITKSLFGSSRPFSGGTSVTVSRNEKVTSYTKLPKGRCSWNNCWLEGLNLPITL 829   
 QY 869 TPKSLYLNLRRPVOAFSADHESPTGDAQARAFKSGHLLNLSPVGVKVDRCSTTHP 928   
 DB 830 SSRILMLKQIIPFVKAEVATYHGGIQTENPBGIRFGHGHLLNVAVPVGVRFKGNHNR 889   
 QY 929 NKYSFMAAICDAYRTISGETITLLSHQETWTTDAFLARHGVVVGRGSMVASLTSNIEVY 988   
 DB 890 DFYTIIVAYAPDVYRNPDCDPTLPINGATWTSIGNNLTRSTLLVQASSHTSVNDVLEIF 949   
 QY 989 GHGREYEDASRGYGLSAGSRVRF 1012



QY	946	SGTETILLSHOETWTTDAPHLARHGTVVVRGSMYASLTSNIEVYGHRYEYDADRGCYGLS	1005
DB	860	PSCTGLAINDVSMILTATNLARQAFIVRAGNHIALTSVGMFQFQFPELSSSRNVD	919
QY	1006	AGSRVRF	1012
DB	920	LGAKVAF	926
RESULT 8			
Q5L6J3 CHLAB			
ID	Q5L6J3	CHLAB PRELIMINARY; PRT; 926 AA.	
AC	Q5L6J3;		
DT	01-FEB-2005	(TREMREL. 29, Created)	
DT	01-FEB-2005	(TREMREL. 29, Last sequence update)	
DT	01-FEB-2005	(TREMREL. 29, Last annotation update)	
DE	Polymorphic outer membrane protein.		
OS	Name=pmp14G; Synonyms=pomp98A; OrderedLocusNames=CAB282;		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=83555;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	STRAIN=S26/3;		
RX	PubMed=15837807; DOI=10.1101/gr.3684805;		
RA	Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,		
RA	Livingstone M., Cerdeno-Tarraga A.-M., Harris B., Doggett J.,		
RA	Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,		
RA	Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.,		
RT	"The Chlamydia abortus genome sequence reveals an array of		
RT	variable proteins that contribute to interspecies variation."		
RL	Genome Res. 15:629-640(2005).		
DR	EMBL; CB848038; CAH63732.1; -; Genomic_DNA.		
KW	Complete proteome.		
SQ	SEQUENCE 926 AA; 98440 MW; F81C2A79A3D575D1 CRC64;		
Query Match 21.3%; Score 1122; DB 2; Length 926;			
Best Local Similarity 30.3%; Pred. No. 1.1e-57;			
Matches 325; Conservative 151; Mismatches 389; Indels 208; Gaps 35;			
QY	1	MOTSHPKPLMLAYSCCSLNGGVAABIMVPO-GIYDGETILTVSPFYTVIGDPSGTVV	59
DB	1	MRPSLYKILISSTLTP-ISPFHSQLHAERVALTOESILDAN-----GAFSPQST	48
QY	60	PSAGELTLKLNLSI-----AALPLSCFGLNLSGTVLGRHSLTPENIRTSNGAA	111
DB	49	STAGG-TIYNVESDISVDVGTAAALASSAFVQTADNLTFKGNHLSLITNAGANPAG	107
QY	112	LNSAADGLFTIEGFKELSPFNSCNLSLLAVLPAATTKGSGTPTTTTSPNGTIYSKTDLL	171
DB	108	INVTADKILTLTDFSKLSFKEC-----PSSLVNTG-----KGAMKGGALN	149
QY	172	LLNNEKPSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGCAQVTVSFSMAN	231
DB	150	LANNASILPDQNYSAENGGAISCKAFSLTGSSEISFTTNSAKKGA-----	197
QY	232	EAPIAFVANVAGVGGGIAAVDQGGVSSSTSTEDPVVPSRNTAVEPDGNVARGVGGI	291
DB	198	-----IAATGIAHLSDNQ-----GTIRFSGNTAV-----NSGGAV	227
QY	292	YSYGNVAFNLNNGKTLFLNNVAPVYIAAKQPTSGOASNTSNNYDGGGAIFCKNGAAGSN	351
DB	228	YSEASMTIAGNNHVAFSNNAVS-----GSSDGGCGGAHCSK-----TG	265
QY	352	NSGVSFDGEGVVFSSNVAAGKGAIAKLSVANCGVQVQLRN-----IANDGGAIVLG	407
DB	266	SAPTITRDNKVLIIPEENTSSAKGAIYTDKILITGSGTAFINNKVTHATPKGAIGIA	325
QY	408	ESGELSADYDGIIPDGNLKRKTAKENAAVNGVTVSSQAISMSGGKLTTLRAKAGHOI	467
DB	326	ANGECSLTAHGDITFDNNLMAT-QDNA-----TIKRNAINTEGKGFVNLRAASGKI	378

```
QY 468 LFNDEPIAMGNQPAQSKLLKINDGEG---YTGDIIVFA-----NGSSSTLY 511
DB 379 SFYDITV-EGV-----AADLLTKAEGDKTYNGRIIFSGEKLTEBQAADVADNKKITFT 432
QY 512 QNVITIEQRIVLREKAKLSVNSLSQTGSL-YMEAGSTWDFVTPQPPQPPAANGLITLS 570
DB 433 QPITLAAGELVLRSGVEAKTVQTAGSLILMDAGT----- 469
QY 571 NLHLSSLLANNVNTPTTPPAQDSHP-AVIGSTTAG-SVTISGPIPFRLDDTAYDR 628
DB 470 ----KLSAKTDATEATNLAINPTLIDGKFAVDAVAAGKQVTLTSGAIGVDPGKPYE- 524
QY 629 YDMLGSNOKIN----VLKQL---CTKPPANAPSDLTIGNMPKYGYQGSWKLAWDPNTA 681
DB 525 -----NHKLNDTLALGGIQLSGKGSVTTTNPVSHV-VGVAEETHYGYGNGNSVSWKDN- 576
QY 682 NNGPYTLKA--TWITGYNPQPERVASLVPNSLMGSIILDIRSAISAIQASVDG-RSYCRG 738
DB 577 NSDPKTOATFTWNTKGYVNPERRAPLVNLSLWGSFIDLRSIQDVLERSVDSILETERG 636
QY 739 LWSGVSNPFYHDDRALQGGVRYISGGYSLGANSYFG-SSMFGLAFTVFGSKDYVVCVR 797
DB 637 LWSGIGNFFFKDORNAERKFRHISGSGYVLGATTTTSRSDLSVAPCQLFAKDXYLVSK 696
QY 798 SNHACIGSVY-----LSTQOALCGSY-----LPGDAPIRASYGFGNQHMKT 839
DB 697 NAANYAGSVYVQHVSKFDDTLRLPNGNWTCCSGFSKSIPIFLDAQI--TYCHTANNMTT 754
QY 840 SYTFAEBSRDWNNCLAGEIGAGLPIVITPSKLYINELRPVQAEFVSADHESPTERGD 899
DB 755 SYTDYPEVKGSGNDTLGLTSTSVPIPVPSISIF-DSYAPFAKLVVYAHQDDPKPTT 813
QY 900 QARAFKSHLLMLSPVGVKPRCSSTHPNKYSFMAVICAAYRTISGTETTLSSHQWTW 959
DB 814 EGRVPESDILLNVSPVIGIKPEKLSYGRSAYDLTMTYIPDYRHNPSCMTGLAINDVSW 873
QY 960 TTDAPHLARHGTVRSGVYSATISNIEVYHGRVYRDASRGYGLSAGSRVRF 1012
DB 874 LTTATNLARQAFIVRAGNHIALTSGVEMPFGFPELRSSRNYNVDLGAQVAP 926
```

## RESULT 9

```
PMP10_CHLPN
ID FMP10_CHLPN STANDARD; PRT; 928 AA.
AC Q9RB65; O86163; Q9RB64; Q986P2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE Probable outer membrane protein pmp10 precursor (Polymorphic membrane
DE protein 10) (Outer membrane protein 5).
GN Name=pmp10; Synonyms=omp5; OrderedLocustNames=CP0303, CpB0467;
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity.";
RL Am. Heart J. 138:8491-8495(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eskin J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
```

```
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Maier R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-914.
RC STRAIN=CWL029/VR-1310;
RX MEDLINE=99081766; PubMed=9864239;
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT "Identification of two novel genes encoding 97- to 99-kilodalton outer
RT membrane proteins of Chlamydia pneumoniae.";
RL Infect. Immun. 67:375-383(1999).
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AJ133034; CAB37071.1; -; Genomic_DNA.
CC EMBL; AB002192; AAF38160.1; -; Genomic_DNA.
CC EMBL; BA000008; BAA98657.1; -; Genomic_DNA.
CC EMBL; AB017158; AAP98398.1; -; Genomic_DNA.
CC EMBL; AJ001311; CAA04671.1; -; Genomic_DNA.
CC PIR; G81591; G81591.
CC PIR; G86546; G86546.
CC PHCI-2DPAGE; O86163; -.
CC TIGR; CP0303; -.
CC InterPro; IPR005546; Auto_transp_beta.
CC InterPro; IPR006315; Auto_transporter.
CC InterPro; IPR011427; ChlamPMP_M.
CC InterPro; IPR003368; Chlamydia_PMP.
CC InterPro; IPR006626; Pdh1.
CC Pfam; PF03797; Auto_transporter; 1.
CC Pfam; PF02415; Chlam_PMP; 7.
CC Pfam; PF07548; ChlamPMP_M; 1.
CC SMART; SM00710; Pdh1; 4.
CC TIGRFAMs; TIGR01414; autotrans_barl; 1.
CC TIGRFAMs; TIGR01376; POMP_repeat; 5.
CC Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 928 Probable outer membrane protein pmp10.
FT CONFLICT 305 305 Missing (in Ref. 4).
SQ SEQUENCE 928 AA; 97230 MW; 0590D5206A1DD0E1 CRC64;
Query Match 20.9%; Score 1101; DB 1; Length 928;
Best Local Similarity 30.5%; Pred. No. 1.9e-56;
Matches 324; Conservative 150; Mismatches 402; Indels 188; Gaps 34;
QY 1 MQTSPHKKFFLSMILA-YSCSLNGCGGVAAMVPGQIYDGETLVSPFYVIGDPSTTV 59
DB 1 MKSQFSLVLSSTLACFTSCSTVFPA-TARNIGPSDFSGSTNVTG--YTPKNTTIGDY 57
QY 60 FSAGELTKNLDNSIAALPLSCFNLGSLFTVLRGHSHTFENIRTSNTGAALNSAAG 119
```

Db	58	TLTQDITLQNLGDS--AALTGKCFSDTTBSLSFAGKGYSLFLNKLSSAEGRAALS-VTTDX	115
Qy	120	LFTIEGFKELSPNCSNLLAVLPAAATTKGSGTPTTTPSPNGTIYSKTDILLNNEKFS	179
Db	116	NLSLTGFSSTLP-----LAAPSSVITTPSG-----KGAVKCGDGLTFDNNGTIL	159
Qy	180	FYSNLVSGDGGDAIDAKSLTVQGISKLCVFOENTAOADGACQVVTFSFAMANEAPFVA	239
Db	160	FQDYCEZENGGAISTNLSKRNSTGISIFEGNKSSA-----	195
Qy	240	NVAGVRGGGIAAVQDQGGVSSSTSTEDPVVFSRNTA-VEPDGNVAR-VGGGIYSYGNV	297
Db	196	--TGKKGGAICA-----TGT-----VDITNNTAPTLSNNIARAAGAINSTGNC	238
Qy	298	AFPLNGKTLFLNNVASPVVIAAKQPTSGQASNTSNNGYGGGAIFCKNGAQAQSNNGSVS	357
Db	239	TTTGNTSLVFSN-----SVTATAGNGGAL-----SGDADVT	270
Qy	358	PDGEGVVPFSSNVAAGKGAIIYAKKLSVAN--CGPVQPLRNI-----ANDGGAIVLBSGG	410
Db	271	ISGNOSVTPSGNOAVANGGAIYAKKLTULASGGGGISFNNIVQGTAGGGAISILAAG	330
Qy	411	ELSLADYGDIIIFDGN-----LKRITAKENAADVNGVTVSSQAISMGGGKIITLRAKAGH	465
Db	331	ECLSAREAGDIFPNGNAIVATTPQITKRNISDI-----GSTAKITNLRAISCH	378
Qy	466	QILFNDPIEWANGNNOPOASSKLLKINDGEG-----YTGDIVPA-----NGS	507
Db	379	SIFPYDPIITA-----NTAADSTDTLNLKADAGNSTDYSGISVFSGEKLSDEAKVADNLT	434
Qy	508	STLYQNTVTEQGRIVLREKAKLSVNSLSOT--GGLYMEAGSTWDFTVTPPQPPQPAANQL	566
Db	435	STLKQPVITLTAGNLVLRGVTLDTKFTGTAGSSVIMDAGTTL-----KASTE	483
Qy	567	ITLSNLHLSSLLANNAVTPNPPNPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAY	626
Db	484	VILTGLSIPVDSL-----GEGKKVVAASAASKNVALSGPILLLLDNQNAV	529
Qy	627	DRYDMLGSKNQKINVLKLO-LGTKPPANAPSDLTLGNEMPKYQGGSKLAWDNTANNGP	685
Db	530	ENHD-LGKTQDPFSFVQLSALGTATTDDPAVPTVATP-THYGYQGTWGMTVDDTAST-P	596
Qy	686	YTLKAT--WTKTYNPGPERRVASLVNSLWGSILDIRSAHSATQASVDGRSYCRGLWVSG	743
Db	587	KTKTATLAWNTGYLPNRPQGPLVNSLWGSFSDIQAIQGVIERGALTLCSDRGPFWAAG	646
Qy	744	VSNFFYHDDRDLGQYRISGGYSL--GANSYFGSSMFLAFTBVFGRSKDYVVCRSNHRA	802
Db	647	VANFLDKKGEKRTKHSKSGGYAIGGAQTCSENILISPAFCQLFGSKDKDPLVAKNHTDT	706
Qy	803	CIGSVYLS-----TQOALCGSYLFGDAPI--RASYGFGNOMKTSYTFBESDV	849
Db	707	YAGAFYIQHITECSGFGICLLDLKPGSWSHKPLVLEGQLAYSHVNDLTKYTAYPEVKG	766
Qy	850	RWNDCLAGBIGAGLPIVITPSKLY-LNELRPFVQABFSDYADHESFTBEGDQARAKSGH	908
Db	767	SWGNNAFNNMLGASSHSY--PEYLHCPTDTYAPYIKMLTYIRQDSFSEKTEGREGSDSN	824
Qy	909	LNLISVPVGVKPFDRCSSTHFNKYSFMAAYICDAVRTISGTEFTLLSHOETWTTDAPHAR	968
Db	825	LFNLSLPIGVKPFKPSDCNDFSDLTLSUYVPDLIRNDPKCTTALVTSGASWETVANNLAR	884
Qy	969	HGVVVRGSMYASLTSTNIEVYGHGRYERDASRGYLSAGSRVRF	1012
Db	885	QALQVRAGSHYAPSPMEFVLGQFVFEVRGSSRIYVNDLGGKQF	928
RESULT 10			
ID	Q823W9	CHLCV PRELIMINARY; PRT; 926 AA.	
AC	Q823W9		
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	







```

DR EMBL; AE001627; AAD18590.1; -; Genomic DNA.
DR EMBL; AE002193; AAP38164.1; -; Genomic DNA.
DR EMBL; BA000008; BAA38654.1; -; Genomic DNA.
DR EMBL; AE017159; AAP98394.1; -; Genomic DNA.
DR PIR; A81591; A81591.
DR PIR; D72078; D72078.
DR PHCI-2DPAGE; Q92393; -.
DR TIGR; CP0307; -.
DR InterPro; IPR005546; Auto_transpbeta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter_1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 930 Probable outer membrane protein pmp8.
FT VARIANT 177 177 T -> A (in strain AR39, strain J138 and strain TW-183).
SQ SEQUENCE 930 AA; 97670 MW; 46A9B5E3BB913C4C CRC64;

Query Match 20.5%; Score 1079; DB 1; Length 930;
Best Local Similarity 28.3%; Pred. No. 3.8e-55;
Matches 303; Conservative 160; Mismatches 407; Indels 202; Gaps 29;

QY 1 MQTSPHKFPLSMILAYSCSLGGGYAAEI-MVPOGIYDGETLTVPSPPTVIGDPSPGTVV 59
DB 1 MKPIPLKLLISLTIV-TPIILLSIATYGADASLSPDSDGAGGSTFTPKST-ADANGTNY 58

QY 60 PSAGELTKUNLDSIAALPLSCFNGLLGSFTVLGRGHSLLTPEINTSTNGAALNSAADG 119
DB 59 VLSGNVYINDAGKG-TALTGCGCTETTGDLTFTGKYSFSPNTVDAGSNAGAAASTTADK 117

QY 120 LPTIEGRKELSFNSCNLLAVLPAATTKNGSQTPPTT---STPSNGTIYSKTDLLLNNE 176
DB 118 ALTFPGPNLSF-----IAAGPTVASGKSTLSSAGALMTDNGTILFQNV----- 164

QY 177 KFSFYNLVSVDGGDAIDAKSLTVQGISKLCVFOENTAQDGAQCVVTSFSMANEAPIA 236
DB 165 -----SNEANNNGGAILTKTSLISGNTSSITFTNSAKLGGALYSSAAASISGNTGQLV 219

QY 237 FVANVAGVRGGIAAVQDQGGVSSSTEDPVVSPSNTAVFDPGNVARVGGGIYSYGN 296
DB 220 FMNN-KGETGGGALGFE-----ASSSITQNSLFFSGNTATD----- 255

QY 297 VAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNVYDGGAIKCKGQAQGSNNSGSV 356
DB 256 -----AA-----GKGGAICYEK-----TGETPTL 274

QY 357 SPDGEVGVVFSNVAAGKGAIAKLVANCGVPVQFLRN-----IANDGGAIVLGSGE 411
DB 275 TISGNKSLTFPENSSTVGGGAIKCHGLDLSAAGPTLFSNNRCGNVTAAGKGAIAIADSGS 334

QY 412 LSLGADYGDIIIFDGNLKRKAKENAADVNGVTVSSQAIMSGSGKITTILRAKAGHQILFND 471
DB 335 LSLSANQGDITFLGN-----TLTSTAPTSTRNAIVLGSASAKITNLRAAQGGQSIYFVD 387

QY 472 PIEMANGN-----NQPASQSKLLKINDGEGYTGDIIPA-----NGSSTLY 511
DB 388 PIASNTTGASDVLITNQPDNSPL-----DYGSTIIVFSGEKLSDRAKAADNFTSILK 440

QY 512 QNVITEQRIVLIREKAKLSVNSLSLTGGSLYMEAGSTWDFVTPPOPPQPPAANOLITLSN 571
DB 441 QPLALASTLAKGNVELDVNGFTQTEGSTLL-----MQPGTKLAKADTAISLTK 490

QY 572 LHLSSLILLANNAVTPNTPNPAQDSHPAVIGSTTAG---SVTISGPIFFEDLDLDYDR 628
DB 491 LVVDLSALEGNKSVS-----IETAGANKTITLSPLVFQDSSGNFYE- 532

QY 629 YDWLGSNQKINVLKQLGTGP-----PANAPSDLTIG-----NEMPKYGYQGSWKLA 675

```

Db 389 ES-----QEDPATNAPATOGGAKEBAANSFDIILKQNVNITFSSNSKTAG 436  
Qy 247 GGIAAQQOQVSSSTEDPVVSFNTAVEFGNVARVGGGTYSGNVAFLNGKTL 306  
Db 437 GAIRV--SGSATIENTGT---CTFSNNNAKE---QGAISVNGNCIDITGNKV 483  
Qy 307 FLNNVASPV-----YIAAKQPTSGOASNTSNNGYDGAIFC----- 342  
Db 484 FSGNOAQRIAPASIVTVBEAAVEEPVAK-----GSGGAIYCVKAPISIPATPIQP 535  
Qy 343 -----XNGAQAQSGNNS-----GSVFDGEGVVPFSSNVAAGKGAIVAKKLS 384  
Db 536 KQTLPSLSAALSGETHAKQBEKSPDPCLTISGNASVIPDNNSSTVTGGAHAKKV 595  
Qy 385 VANCQPVQFLRIANDGGAIYLBESGELSLSDYDILFDGNLKEKAKEN----- 434  
Db 596 LSSNGMTFSSNNSGKGAIYADGDISITATTGSIIPQGN-KVTAADSIPLTKKEIA 654  
Qy 435 --AADVNGVTVSSQA-----ISMGSGGK 455  
Db 655 IAAESIEBKTFPSQASGSASAGTSAFTFLANKAEIPAESQAKENSKPTCSNHLGSGAK 714  
Qy 456 ITTLRAKAGHOILFNDPI-----EMANGNN-----QPAQS 485  
Db 715 ISQLRAQTGQTFFVDPIITTAPAAAUVVTAQPEASLAKATSGIPASASVSPAPAPT 774  
Qy 486 SKLLKIN-----DGEYTGDIYPA-----NGSSTLYQNV 514  
Db 775 KTLKINAPDQDPEIQKVAEAAQOASVYNGKIVFSGEKLSEDAKNPLNATSVIHNDV 834  
Qy 515 TTEQGRVILREKALSVNSLSOTGSLVMEAGSTWDF--VTP-----QPPQOPPA----- 562  
Db 835 SLEAGTLVSSGAGLLVDSTQKEGSLIVDGGTSIITNTVASEGLQSRSTPPSPKNAI 894  
Qy 563 -----ANQL-----ITLSNHLSSLLANNVTVN 587  
Db 895 PVIRAVSKVIASSLINLERADSGAGAVVPTIEESPDGSIITNLAVNLDLSLENGKVIT- 953  
Qy 588 PPTNPAQSHPAVIGSTTAGSVTISGPIFFEDLDATADYRDVWLGSQKINVLKQLGT 647  
Db 954 -----LAAKGGSGSVTLTGDLQFQDSSQNFYDN-PLLNKNFTLNFLDI---S 996  
Qy 648 KPPAN-----AFSDTLGNEMPKYGVQSGKWLADPNTANNPPTLKATWTKTGVN 698  
Db 997 APDAEKIHTEGFIIPQAGTSSN-----LGQGGKWEVTEVDKSGKVSFEKK--WVSTGYI 1050  
Qy 699 PGPERSVASLPNSLWGLSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLGQ 758  
Db 1051 PTANRATLVPSVWCSAIDMRAFNQNLVEVSTEGEDPHRGLWISGISNFFHKDSTKVQEG 1110  
Qy 759 YRISGYSLGANSY-FGSGMFLGAFTEVFGRKDVVVCNSNHACIGSVYLSLQALCG 817  
Db 1111 FRHISGYSVGVSTQPTSNKVMDLAFQMLGSKDYRLADARSHVYAASHTTKCEK-LVN 1169  
Qy 818 SYLFG-----DAFIRSYGFGNQHMKTSTYFPAESDVWNNCLAGE 859  
Db 1170 HYTFSKRKGAIALARKPEKSPIIFDA--QLSYSLSHNSMTTKTTPNPSRSSGNNHCVAGE 1227  
Qy 860 IGAGLPITVIT-PSKLYNLNLRPFVQAEFSYADHESF--TEEGQARAFKSGHLNLSVPV 916  
Db 1228 LGSYLPILVDHPA---IEELFPFKLHVIVFQEDDFKETOGGTENRFQSAHPVNSLPL 1284  
Qy 917 GVFPDRCSSTHPNKSYPMAIYICDAYTISGTETLLSHOETWTTDAFLHARGVVVRGS 976  
Db 1285 GVRFEKTNKL--NTYNIRLAYQPDYRDAPKSVFLPSVHTAWSTGTATNLRSQAMILDGS 1342  
Qy 977 MYASLTNSIEVYGHGRYEYRDASRGYGLSAGSRVP 1012  
Db 1343 DHHHLTDNLEVFCHGAFELRGSSRNYNVDIGGRYKP 1378

RESULT 14  
PMP7\_CHLPN

ID AC PMP7\_CHLPN STANDARD; PRT; 936 AA.  
DT Q92986; Q9JS42; Q925Q4;  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Probable outer membrane protein pmp7 precursor (Polymorphic membrane protein 7) (Outer membrane protein 12).  
GN Name=pmp7; Synonyms=omp12; OrderedLocNames=CpN0445, Cp0308, CpB0462;  
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=CWL029;  
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;  
RA Kalman S., Mitchell W.P., Marathe R.W., Lammell C.J., Fan J., Hyman R.W.,  
OLinger L., Grimwood J., Davis R.W., Stephens R.S.;  
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."; Nat. Genet. 21:385-389 (1999).  
RL [2]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=AR39;  
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.P.,  
White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,  
Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,  
Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,  
McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39."; Nucleic Acids Res. 28:1397-1406 (2000).  
RL [3]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=J138;  
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;  
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA."; Nucleic Acids Res. 28:2311-2314 (2000).  
RL [4]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=TW-183;  
RX Geng M.M., Schumacher A., Muehlhofer I., Bensch K.W., Schaefer K.P.,  
Schneider S., Pohl T., Essig A., Marre R., Melchers K.;  
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis."; Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.  
RL [5]  
RN NUCLEOTIDE SEQUENCE OF 658-936.  
RC STRAIN=CWL029 / VR1310;  
RX MEDLINE=20007584; PubMed=10539856;  
RA Christensen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,  
Madsen A.S., Knudsen K., Falk E., Birkelund S.;  
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity."; Am. Heart J. 138:S491-S495 (1999).  
RL -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).  
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.  
CC -1- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC  
CC ENBL; A8001627; AAD18589.1; -; Genomic DNA.  
CC ENBL; A8002193; AAF38165.1; -; Genomic DNA.  
CC ENBL; BA000008; BAA98653.1; -; Genomic DNA.  
CC ENBL; AB017158; AAP98393.1; -; Genomic DNA.  
CC ENBL; AJ133034; CAB37067.1; -; Genomic DNA.

```

DR PIR; B81591; B81591.
DR PIR; C72078; C72078.
DR PIR; C86546; C86546.
DR PHCI-2DPAGE; Q92898; -.
DR TIGR; CP0308; -.
DR InterPro; IPR005546; Auto transportheta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam PMP; 7.
DR Pfam; PF07548; ChlamPMP M; 1.
DR TIGR; TIGR01414; autotrans bar1; 1.
DR TIGR; TIGR01376; POMP_repeat; 5.
DR Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 23 Potential.
FT CHAIN 24 936 Probable outer membrane protein pmp7.
FT CONFLICT 658 666 PTRHGFRHI -> EDNIYRHN (in Ref. 5).
FT CONFLICT 822 822 Y -> H (in Ref. 1, 4 and 5).
SQ SEQUENCE 936 AA; 100107 MW; 3981DB3C950AF95A CRC64;

Query Match 19.7%; Score 1040; DB 1; Length 936;
Best Local Similarity 29.0%; Pred. No. 7.8e-53;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

QY 1 MQTSPHKFPLSNILAYSCSLNGGYYAABWV--PQGIYDGETLTVSPYTVIGDPGOTT 58
DB 1 MKSSVSWLFFS8IPLF8LSI-----VAEVLDSNNSYDGSNGTTFVTFTDAAAGTT 56

QY 59 VFSAGELTKNLDNSIAALPLSCFNLLGSPVTLQRGHSLTFENIRT-STNGAALSNSAA 117
DB 57 YSLSDV8FQAGALGIPLAGCFLEAGCDITFQGNQHALPAPINAGSSAGTVA8TSA 116

QY 118 DGLFTIEGKELSFNSCNLSLAVLPAATNNKSGOPTTTTSPNGTIYSKTDLLLNNEK 177
DB 117 DKNLLFNDFSLSIISCPSLLL-----SPT-----GQCALSKVGNLSLTGNSQ 159

QY 178 FSPY8NLYSGDGAIDAKSLTVQIGSKLCVQENTAQADGACQVTVSFSAMANEAP 237
DB 160 IIFTQNFSD8G8VINTKFLNLLSGTSPAF8SRNQA----- 195

QY 238 VANVAGV8GGIAAQDQCGQV88STSTEDP-VVSP8RNTAV8P8GNV8GV8GIYSXN 296
DB 196 ---FTGKGGV8YA-----TGTTIENS8QIV8FSQNL8-KGS-----GGALYSTDN 238

QY 297 VAPLNGKTLFLNNV8PVYIAAKOPT8GQASNT8NNTY8DGGAI8FCK8GAQ8GNS8GV 356
DB 239 CSITDN8QVIFD8NSA-----WEAAQ-----AQQ8AICC-----ITTDKTV 274

QY 357 SP8G8GV8F88NV8AG8G8I8YAK8L8V8ANG8PV8Q8FL8NT8ND-----GG8I8YL8ES8G 410
DB 275 TLTG8KN8L8PT8NT8AL8TY8G8A8IS8L8K8V8IS8AG8PT8L8F8Q8S8T8SG8S8AG8GG8G8AIN8IAS8G 334

QY 411 ELSL8ADY8GDI8FDGN8L8K8RTAK8NA8DV8NTV8SQ8AIS8MG8SG8K8IT8L8RA8K8GH8L8FN 470
DB 335 EL8ALS8AT8G8D8IT8F8NN-----QVT8NG8ST8R8N8I8N8IT8AK8V8IR8AT8Q8SI8FY 385

QY 471 DPI8M8ANN8P8AQ8SK8LLK8IND8EG-----YT8GD8IV8-----ANG8S8T8LY8Q 512
DB 386 DPI8T-----NP8GTA8ST8DT8L8NL8AD8AN8SE8I8Y8G8AI8V8F8G8K8L8P8TE8KA8IA8AN8VT8T8IR8Q 441

QY 513 NVT8IE8Q8R8IV8RE8K8AK8L8V8NS8L8SQ8GS-LY8WE8AG8ST8W8D8FT8PP8Q8PP8A8N8Q8L8IT8SN 571
DB 442 PAV8L8ARG8DL8V8G8VT8FT8K8D8L8T8Q8PS8GR8I8L8MD8G8T8L8S-----AK8E8AN8LS8I8NG 490

QY 572 LHL8LS8SL8L8ANN8V8T8P8T8P8AQ8D8HP8AV8IG8T8TAG8V8T8SG8T8T8AG8T8D8D8T8AY8D8V8M 631
DB 491 LAV8L8SS8LD8G8TN8K8-----ALK8TE8AD8KN8IS8LS8GT8L8I8D8T8EG8SF8Y8HN8- 535

QY 632 LGS8NQ8K8I8NV8L8QL8GT8K8PP8N-----APS8DL8TL8GN8P8K8Y8G8Q8SK8W8K8L8AD8P8NT8ANG8P 685
DB 536 LKS8AST8Y8L8LE8L8-----TT8AG8NT8IT8L8G8L8ST8LT8L8Q8EP8TH8Y8G8Y8Q8N8QL8SW8AN8T8SS8K-- 590

```

"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CML029 from USA."; [5]  
 Nucleic Acids Res. 28:2311-2314 (2000).  
 [5]  
 NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 STRAIN=TW-183;  
 Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,  
 RA Schneider S., Pohl T., Essig A., Warre R., Melchers K.;  
 RA "The genome sequence of Chlamydia pneumoniae TW183 and comparison with  
 RT other Chlamydia strains based on whole genome sequence analysis."; [5]  
 Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RL  
 CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)  
 CC (Potential).  
 CC  
 CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.  
 CC  
 CC -!- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL: AJ133034; CAB37069.1; -; Genomic DNA.  
 DR  
 CC EMBL: A2001628; AAD18591.1; -; Genomic DNA.  
 DR  
 CC EMBL: A2002192; AAP38163.1; -; Genomic DNA.  
 DR  
 CC EMBL: BA000008; BAA98655.1; -; Genomic DNA.  
 DR  
 CC EMBL: AB017158; AAP98395.1; -; Genomic DNA.  
 DR  
 CC PIR: B72077; B72077.  
 DR  
 CC PIR: B86546; B86546.  
 DR  
 CC TIGR: CP0306; -.  
 DR  
 CC InterPro: IPR005546; Auto transportbeta.  
 DR  
 CC InterPro: IPR006315; Autotransporter.  
 DR  
 CC InterPro: IPR011427; ChlamPMP\_M.  
 DR  
 CC InterPro: IPR003368; Chlamydia\_PMP.  
 DR  
 CC InterPro: IPR006626; PbhA.  
 DR  
 CC Pfam: PF03797; Autotransporter; 1.  
 DR  
 CC Pfam: PF02415; ChlamPMP; 7.  
 DR  
 CC Pfam: PF07548; ChlamPMP\_M; 1.  
 DR  
 CC SMART: SM00710; Pbh1; 5.  
 DR  
 CC TIGRFAMs: TIGR01414; autotrans bar1; 1.  
 DR  
 CC TIGRFAMs: TIGR01376; POMP repeat; 6.  
 DR  
 CC Complete proteome; Membrane; Multigene family; Outer membrane; Signal.  
 KW  
 FT SIGNAL 1 26 Potential.  
 FT CHAIN 27 928 Probable outer membrane protein pmp9.  
 FT SEQUENCE 928 AA; 98333 MW; 58910A8F04F12219 CRC64;  
 SQ  
 Query Match 19.4%; Score 1021; DB 1; Length 928;  
 Best Local Similarity 29.5%; Pred. No. 1e-51;  
 Matches 313; Conservative 161; Mismatches 404; Indels 184; Gaps 39;  
 QY 1 MOTSFHKPFLSMILAYSCCSLNGGVYAEIMV---PQGIYDGE-TLTVSPYTVIGDPDG 56  
 DB 1 MKSSLHFLFISSSLALP-LSLNFSAFAVVEINLPTNSPFGCTYV---PPAQTTNADG 56  
 QY 57 TTVPFAGELTKLNDLSIALPLSCFNLGLSPVLGRHSLTFENRTSTNGAALSNSA 116  
 DB 57 TIYNLTGDVSRITNA-GSFALTALSCPKETTNLSLFGQHGQYQLQNDAGAN-CTFTNTA 114  
 QY 117 ADGLFTIEGPKELSFNSCNLSLAVLPAATTNKGSTPTTTSTPSNGTIYKTDLLILNNE 176  
 DB 115 ANKLLSPSGFYSL- - - - - IQTWATGTGAIKSTGACSIQSNY 154  
 QY 177 KFSFYSNLVSDGDGAIDAKSLITVQGISKLCVQBNTAQADGGACQVVTFSAMANEAPIA 236  
 DB 155 SCYFGQNFNDGALGSSLSL-SLNPNTFAKNKATQKGGAL- - - - -YST- - - - - 200  
 QY 237 FVANVAGVGGIAAVODGQGVSSSTSTEDPVVPSRNTAVEFDGNVARGGIIYSGN 296  
 DB 201 - - - - -GGIT- - - - -INNLTNS- - - - -ASPSENTA- - - - -ANNGGAIYTEAS 231  
 QY 297 VAFLLNGKTL-FLANVASPVYIAAKQPTSGQASNTSNNGYDGGAIKCKNGAQAQSNNSGS 355  
 DB 232 -SFISNNKAISSFINN- - - - -SVTATSAT- - - - -GGAIYC- - - - -SSTSAPKPV 268

356 VSPDEGVVFPSSNVAAGKGAIIYAKKLSVANCGPVPQFLRNAND- - - - -GGAIYLGSSG 410  
 DB 269 LTLSDNGELNFIGNTAITSSGGAIVTDNLVLSGGGPTLPKNSAIDTAAPLGGAIADSG 328  
 QY 411 ELSLSADYVDIIFDGNLKRKAKENADVGVTVSSQALSMG-SGGKITTLAKAGHQILP 469  
 DB 329 SLSLSALGDDITFEGN- - -TVVKGAS- - -SSQTTTRNSINIGNTNAKIVOLRASQGNITLYP 383  
 QY 470 NDPIEMANGNNOQAQSSKLLKINDGE- - -GYTGDIVPA- - - - -NGSSTIY 511  
 DB 384 YDPIT- - -TSITALSALMLNGPDLACNPAYQGTIVFSGEKLSEAEAEADNLKSTIQ 439  
 QY 512 QNVTIEQRIVLREKAKLSVNSLSQTGS-LYMEAGSTWDFVTPPOPPPPAANOLITLS 570  
 DB 440 QPLTLAGQLSLKSGVTIVAKSPSQSPGSTLLMDAGTTLETADG- - - - -ITIN 487  
 QY 571 NULHLSLSLANNVNPPTPPAQDSHPANVIGSTTAG-SVTISGPIPPEDLDLDDTAYDRY 629  
 DB 488 NLVLNVDLSL- - - - -KETKATLKATQASQVTVTLGSLSLVDPSPGNVYEDV 532  
 QY 630 DWLGSNOKINVLKLOLGTKPPANAP-SDLT- - - - -LGNEMPKYGYQGSWKLWDPNTANNGP 685  
 DB 533 SW- - -NNPQV-FSCLTLTADDDPANIHITLADPLKPNPHWGYQGNWALSQWEDATYKSK 589  
 QY 686 YTLKATWTKTGYNPGPVERVASLVPNSLMGSIILDIRSAHSAIQASVDGRSYCRGLWVGVS 745  
 DB 590 -NATLWTWTKYGNPNPERRGTLVANTLMGSPVDVRSIQQLVATKVRSQOETRGWCEGIS 648  
 QY 746 NFFYHRRDALGQGYRIYSGVYSLGANSYFGS-SMFGLAFTVFGRSKDYVVCRSNHHACI 804  
 DB 649 NFFHKDSTKINKGPRHISAGYVVGATTTLASDNLITAAFCQLFGKDRDFINKNRASAYA 708  
 QY 805 GSVYLSLTOALCG- - - - -SVLFG- - - - -DAFIRASYGFGNQMKTSTYFAEEDSVRW 851  
 DB 709 ASLHLQHLATLSSPSLLRYLPGESEQVLPFAQI- - -SVIYKNTMKTYYTOAPKGSNW 766  
 QY 852 DNNCLAGEIGAGLPITVITPSKLYLNLRLPFOAEFSYADHESFTEBGDQ-ARAFKSGHLL 910  
 DB 767 YNDGCALELASSLPHTALSHEGLFAYFPFIKVEASYIHQDSFKERNTTLVRSFDSGLDI 826  
 QY 911 NLSVPVGVKDFRCSSTHPNKYSFMAAICDAVRTISGTETTLTSHOETWTTDAPHARHG 970  
 DB 827 NVSPVIGITFFERFSRNERASYEATVIYVADVVRKPNPDCCTALLINNNTSKTGTNLSRQA 886  
 QY 971 VVVRGSNYASLSTNIEVYGHGYEYRDASRGYGLSAGSRVP 1012  
 DB 887 GIGRAGIFYAFSPNLEVTNLSMEINGSSRSYNADIGGRKQF 928  
 RESULT 16  
 PMP6\_CHLPN STANDARD; PRT; 1276 AA.  
 ID AC Q92899; Q9JRW2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Probable outer membrane protein pmp6 precursor (Polymorphic membrane  
 protein 6).  
 DE Name=pmp6; OrderedLocNames=CpN0444, CP0309, CpB0460;  
 GN Chlamydia pneumoniae (Chlamydophila pneumoniae).  
 OS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.  
 OC NCBI\_TaxID=83558;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=CWL029;  
 RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;  
 RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."; [2]  
 RL Nat. Genet. 21:385-389 (1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].



```
QY 846 ESDVNDNCLAGEIAGLPIVITPSKLYLNLRLPVPQAEFSAHESFTEEGDOARFK 905
DB 1113 QGKSDWDSPFAVBVGSLFVDL--NRYLTLSYSPVYKLQVVSQKQFQVAAFRIPD 1170
QY 906 SGHLNLSPVGVKDRCSSTHPNKYSFMAAICDAYRTISGTETLLSHQETTTTDAFH 965
DB 1171 ASHLNVNVPGLTTPKESAKPSALLLTGLVAVDAYRDHPHCLTS-LTNGTSMSTFATN 1229
QY 966 LARHGVVVRGSMYASLTSNIYVGHGRYBYRDAASRGYGLSAGSRVP 1012
DB 1230 LSKQAPFAEASGHLKLLHGLDCFASGCELSRSSRYNANCGRYSF 1276

RESULT 17
Q823X0 CHLVC
ID Q823X0 CHLVC PRELIMINARY; PRT; 866 AA.
AC Q823X0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family
DE protein/autotransporter.
GN OrderedLocusNames=CCA00283;
OS Chlamydia caviae.
OS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapfel E.K., Khouri H.M., Federova N.B.,
RA White H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147(2003).
RL EMBL; AS016995; AAP05034.1; -; Genomic_DNA.
DR TIGR; CCA00283; -;
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005546; Auto transptbeta.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP; 1.
DR Pfam; PF02415; Chlam_PMP; 5.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 4.
KW Complete proteome.
SQ SEQUENCE 866 AA; 92969 MW; 6CPA14937C69DB30 CRC64;

Query Match 19.1%; Score 1008; DB 2; Length 866;
Best Local Similarity 28.2%; Pred. No. 5.5e-51;
Matches 303; Conservative 144; Mismatches 358; Indels 268; Gaps 30;

QY 1 MQTSFHKFPFLSMILAYSCCSLNGGYAAEIMVPGIYDGETLVSP-----PVTVI 51
DB 1 MKHPVYVFWLVSSGLAS-TLSLFAAAVQETLNSGDSYNGNATTDAFQTKETTTGAEYCE 59
QY 52 GDSPTTVFSGAGELTLKNLNSIAALPLSCFNLGLSFTVLGRHSLTFENIRTSNGAA 111
DB 60 GN---VCITVAGKT-----PLKSCFETETENLTFIGQYSLCFDNITATAPAA 107
QY 112 LNSAAGDLFTIEGFKELSPNCSNLLAVLPAATNKGSOFTPTTSPSNGTIYSKTDLL 171
DB 108 IEVSAADKLTLSISGSLFSCSYC-----PPGTTGGAKSKSGGTAT----- 147
QY 172 LNNKEFSFYNLVSGGGAIDAKSLTVQGISKLCVQENTAOAGGACQVWTSFSAMAN 231
```

```
DB 148 FDNNSKLLFKNSCSSEGGAINCKSLTLKN-SSVC----- 181
QY 232 EAPIAFVANAVGRGGIAAVQDQGOQGVSSSTSTEDPVVFSRNTAVBFQGNVARVGGI 291
DB 182 -----ANFKNSDK----- 191
QY 292 YSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGOASNTSNNGYGGGAIFCKNGAQAASN 351
DB 192 -----KGGAIYC----- 198
QY 352 NSGVSVPDGEVVPFSSNVAAGKGAIIYAKKLSVANCSPVQPLRNIA-----DGGAILV 406
DB 199 SDENLQLENQMLFSENTSKEEGAIYAKKLSIISGGTFLFSNNTSKAADPKGGAICI 258
QY 407 GES-CELSISADYGDIIIFDGNLKRITAKENAADVNGVTSSOASISMGSGGKIITLAKAGH 465
DB 259 ADADSECSLITANGDIIIFDGNKIITGTPTSKRN-----SIDLGGGKFSQLARADGF 311
QY 466 QILFNDPIEMANGNNQPAQSSKLLKINDGEG---YTGDIVFA-----NGSST 509
DB 312 GVFFYDPI-----ANN--GSDTDTLEINKADGAATYSGRIVFSGEKLTEDEKQVTDNLKSF 365
QY 510 LYQNTIISQGRIVLREKAKLSVNSLSQTGSSLYMEAGSTWDFVTPQPPQPPAANQLITL 569
DB 366 FKQPLTVSGSGSVLKNGVTVSAKQITQSGGAIEMDAGTN-----LTS 407
QY 570 SNLHLISLSLLANNAVTPNPPPAQDSHPAIVGSTTAGSVTISGPIPPEDLDDTAYDRY 629
DB 408 TTEDISLSNLVINTASLGGGVPLAAQ-----ISAGTNKSVTISS-LNLVDADGNGYEP 462
QY 630 DWLGSNQKINVLKQL--GTPKANAPSDITLGNEMPKYGYQGSWKLAMPDNTANNPYTL 688
DB 463 VFTTREPSSIIEAKANGTGTPTIPTHLDHAPAAHYGYQLWTTSWAQGTATTSQLAT 522
QY 689 KATWTKYGNPQPERVASLVPNSLWGSILDIISAHSAIOASVDGRSYCRGLWVSGVSNPF 748
DB 523 LA-WQOTGYNPNPQGPVPLVNTLWGSFSDVRAIQNLMDISVNGADYQRLWASGLANFL 581
QY 749 YHNRDALGOYRYISGGYSLGANS-YFGSSMFGLAFTVEFGRSKDVVCRSHHHCIGSV 807
DB 582 QKSGTETKRRKFRHHSAGYVILGAYAKTLSDDVFSAPFCQLFGDKDLVSKNSNIYAGSI 641
QY 808 YLSTQOALCGSYLFGDAP-----IRASYGFGNQHMKTSYT--FAEESD 848
DB 642 YQHTS-----FWDADNLQLSTLGAQAPLVLAQLTVSHTSNDMKNMTTKYAPQGV 694
QY 849 V-----RWDNNCLAGEIAGLPIVITPSKLYLNLRLPVPQAEFSAHESFTEEGD-QA 901
DB 695 VYPEIKGDWGNDCFGVELGATVPIESPYSLP-DMYSPLRFQLVYAHQEDPKENNSTEG 753
QY 902 RAPKSGHLNLSPVGVKDRCSSTHPNKYSFMAAICDAYRTISGTETLLSHQET--W 959
DB 754 RYFESSDLTNLSMPICGVKPERFSDNDIASYNTVLAAPDLVRSNPPCKTSLVSPITAVW 813
QY 960 TTDAFHLARHVVVRGSMYASLTSNIYVGHGRYBYRDAASRGYGLSAGSRVP 1012
DB 814 LTKATNLARHAFIVKAGNYLSLSNFEIIFSQGFEFGSSRTYNDVLSGSKIQF 866

RESULT 18
Q9RB71 CHLPN
ID Q9RB71 CHLPN PRELIMINARY; PRT; 772 AA.
AC Q9RB71; Q7BX22;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Pmp_3 (Outer membrane protein 5).
DE Names=pmp_3_2; OrderedLocusNames=CpB00018;
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN NUCLEOTIDE SEQUENCE.
```



```

RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and WOL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TW-183;
RA Geng M.M., Schummacher A., Muehlhofer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BA009008; BAA98226.1; -; Genomic DNA.
DR EMBL; AE017157; AAP97951.1; -; Genomic DNA.
DR PIR; H86492; H86492.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005546; Auto_transptbeta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter_1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR TIGRFAMs; TIGR01414; autotrans_barl; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
DR TIGRFAMs; TIGR01376; POMP_repeat; 3.
SQ SEQUENCE 772 AA; 82931 MW; 484FC56D635801EB CRC64;

Query Match 18.6%; Score 978.5; DB 2; Length 772;
Best Local Similarity 29.4%; Pred. No. 2.6e-49;
Matches 260; Conservative 142; Mismatches 316; Indels 167; Gaps 26;

QY 180 FVSNLVSGDAIDAKSLTVQIGSKLVFOENTAGCGACQVVTSPSAMANEAPIAFA 239
DB 3 FPKFSTDNGAINTAKTSLTGTMSALFSENSSKGGAIQTSDALTITGNQGEVSP-- 60

QY 240 NVAGVRGGIAAQQDQGGVSSSTSTEDPVVFSRNTAEPFGNVARVGGGIYGVNAP 299
DB 61 -----SDNTSSDS-----GAAIFTEASVTI 80

QY 300 LNNKTLPLNNVAPVYIAAQPTSGQASNTSNNGDGAIFCKNGAAGNNSGSVDP 359
DB 81 SNNKVSFDN-----KVTGASSTTGDW-SGGAICAYK-----TSTDTKVTLT 123

QY 360 GEGVFFSNAAGKGAIAKLSVANGCPVQFLNAN-----DGGAIYLGSGGELS 414
DB 124 GNQMLFENNTSTAGGAIYVKLELGGITLFSRNSVNGGTAPKGAIAIEDSGELS 183

QY 415 SADYGDIIFDGNLKRITAKENADVNTVSSQAISMGGKITTILRAKAGHQIILFNDPI 474
DB 184 SADSGDIVLGN---TVTSTTPGNT---RSSIDLGTSAKWTALRSAGRAIYFDPI 235

QY 475 MANGNQPQASQSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNTVI 516
DB 236 TGSSTT-----VTDVLKNETPADSALQYTGNIIFTEKLSSETAADSKNLTKLQPVTL 291

QY 517 EQGRVILREKAKLSVNSLSQTGG--LYMEAGSTWDFVTPQPPQPPAANQLITLSNHL 575
DB 292 SGGTSLKHGVTLTQATQADRLNDVGTTL-----PADTS---TINNLVIN 339

QY 576 LSSLANNAVNTNPPPAQSDHPAVIGS--TAGSVTISGPIFFEDLDDTAYDRYDMLGS 634
DB 340 ISSI-----DGAKKAKIETKATSKNLTLSGTITLDDPTGTYENHS-LRN 383

QY 635 NQKINVLKQL-QYKPPANASDITLGNEMPKYQGSW-KLAWDPNTANNGPYTLKATW 692
DB 384 PQSYDILELKASGVTSTAVTDPDTMGKFP-HYGYQGTMGPIVNGTGASTATP-----NW 438

QY 693 TKTYNPGPERRVASLVNLSMGSITLDIRSAHSATQASVDGRSYCRKGLWVGVSVPFYH 752
DB 439 TKTYINPERIGSLVPSNLNNAFIDISLHYLMETANEGLQGDRAFWCAGLSNFFHKDS 498

```

```

QY 753 DALCGQVRYISGGYSLGANSYFGS-SMFGLAFTVEVFGSRKDYVYVCRSNHHACISGVYLS 811
DB 499 TKTRGRFHLISGGVIGGNLHTCSDKILSAAPCOLFGRDRDYFAKNG-----GTVYGGT 553
QY 812 QOALCGSVLFGDAFI-----RASYGFGNQHMKTSTYFAES 847
DB 554 LY-----YQNETYISLPCKLRPCSLSYVPTFIPVLFSGNLSTYTHTDNDLTKYTYPTV 608
QY 848 DVRMNNCLAGEIGAGLPIVITPSKLYLNELRPFVQAEFSYADHESPTTEGDQARAPKSG 907
DB 609 KGSWGNDSFALEFGGRAPICLDESALP-EQWMPMKLQFVYAHQEGKEQTEAREFGSS 667
QY 908 HLNLNLVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFLA 967
DB 668 RLNLALPIGIRPKESDCQDATYNTLGYTVDLVRSNPDCTTTTLRISGDSMKTFGTNLA 727
QY 968 RHGVVVRGSMYASLTNSIEVYHGRIYRDSRSGYGLSAGSRVRF 1012
DB 728 QOALVLRAGNHFCFNSNFEAFSPQSFELRGSSRNYNVDLGAQYQF 772

RESULT 19
Q822Q5 CHLCV PRELIMINARY; PRT; 841 AA.
AC Q822Q5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family
DE protein/autotransporter.
DE OrderedLocustNames=CCA00624;
OS Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OC NCBI_Taxid=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapple E.K., Khouri H.M., Pederoza N.B.,
RA Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae.";
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016996; AAP05366.1; -; Genomic DNA.
DR TIGR; CCA00624; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005546; Auto_transptbeta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter_1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 4.
DR TIGRFAMs; TIGR01414; autotrans_barl; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 3.
DR Complete proteome.
SQ SEQUENCE 841 AA; 90411 MW; 21A9ED16E0D7C65D CRC64;

Query Match 18.4%; Score 971.5; DB 2; Length 841;
Best Local Similarity 27.7%; Pred. No. 7.6e-49;
Matches 298; Conservative 132; Mismatches 347; Indels 299; Gaps 30;

QY 1 MQTSFHKKFFLSMLAYSCCSLNGGYAA---EIMVPOGIYDGETLTVSF-----PY 48
DB 1 MKHPVYVFLVSSGLLASTSS---SPAAAVOETLNSDSYNGNTTTTTFVFPKSTGAEY 56
QY 49 TVTGDPDSGTVFSAAGELTLKLNLSIAALPLSCFGLNGSFTVLGRGHSITPFIKRTSN 108

```

Db 57 TCNGN---VCITYAGKTT-----PLTKSCFTETTENLFLGQGYSLCFDNIATAK 104  
Qy 109 GAALNSAADGLFTTEGKELSFNCNSLLAVLPAATTKNGSQTFITTSNGIYIYKT 168  
Db 105 PAATNSAADKTLISGFSFLPSCSYC-----PFGITTEQ----- 137  
Qy 169 DLLLNNKPFSPYSLNLSGDGAIDAKSLTVQGISKLCVFPQENTAQADGACQVVTSPSA 228  
Db 138 -----CAIQSGKV----- 146  
Qy 229 MANEAPIAFVANVAVRGGIAAQQGQGVSSSTSTEDPVVFSRNTAVEFDGNVARVG 288  
Db 147 ----- 146  
Qy 289 GGIYSYGNVAFLLNGKTLFLNNVAVPVYIAAKQPTSGQASNTSNYDGGAIFFCKNGQA 348  
Db 147 -----TFSNNDKIIFPKNCSTE-----XGGAIKCDTGNA 176  
Qy 349 GSNNSGVSFDEGVVFPSSNVAAGKGAIAKLSVANGCPVQFLRNIA-----DGGA 403  
Db 177 -----ELKPEGNKYLFPSCNSQEGGAIYAKLSIISGGTLPFSNNTSKAADPKGA 230  
Qy 404 IYLGES-GBLSADYGDIIIFGNLKRKTAKENADVGVTVSSQAISMGSGGKITTLRAK 462  
Db 231 ICIAADASECSLTAENGDIIFDGNKIITGTPTKEN-----SIDLGGGKFSQLRAR 283  
Qy 463 AGHQLFNDPIEMANGNNOPOASSKLLKINDGEG---YTGDIIVPA-----NG 506  
Db 284 DGFQGVFFYDPI-----ANN--GSDTDTLEINKADGAATYSGRIVFSGEKLTEDEKQVTDNL 337  
Qy 507 SSTLQNVITIEGRIVLREKAKLSVNSLSQTSGLSIMEAGSTWDFVTPPQQPPAANQL 566  
Db 338 KSFFQKPLTVGGSGFVLKNGVTVSAKQITQSGAIENDAGTN----- 379  
Qy 567 ITLSNLHLISLSLANNVNTPTNPAPQDSHPAVIGSTTAGSVTISGPIFFEDLDITAY 626  
Db 380 LSTTEDIISNLVINTASLGSGGVPLAAQ-----ISAEGTKSVTISL-LNLVDADNGY 434  
Qy 627 DRYDMLGSNOKINVLKQL-GTKPPANASDLTLGNEMPKYGVQSGKWLAWDNTANNCP 685  
Db 435 EYVPFSTTRFPFSLIEAKANGTGTPTIPTTLTHDAPAAHYGVGLWTTSWAQGTATTSSQ 494  
Qy 686 YTLKATWTKYNGPGRVAVSLVPSNLGSLIDIRSAHSAIOASVDGRVYCRGLWVGVS 745  
Db 495 LATLA-WQOTGYNPNPERQGPLVNTLWGSFSDVRAIQNLMDISVNGADYQRLWASGLA 553  
Qy 746 NPFYHDDALGQYRYISGYSIGANS-YFGSMFGLAFTEVGRSGKVVVCRSNHACI 804  
Db 554 NFIQKSGTETKFRHRSAGVVLGAYAKTLLSDVPSAACPQLGRDKDVLVSKNSNIYA 613  
Qy 805 GSYLSTQQAALCGSYLFGDAP-----TRASYGFGNQHMKTSYT--PAE 845  
Db 614 GSIYYQHTS-----FMDAUNDLLQSTLGAQAPLVLAQLTYSHTSNDKMTNMTTKYAP 666  
Qy 846 ESDV-----RWDNCLAGHIGAGLIVITPFSKLYLNLRLPFPVQAQSFYADHESFTEGD 899  
Db 667 QGVYYPEIKGDWDCGFCVGLGATVPIESPSLSF-DMYSFPLRFQLVVAHQDFFKENNS 725  
Qy 900 -QARAFKSHLLMLVVPVGVKFCRCSSTHBNKYSFMAAYICDAYRTISGTETLLSHQET 958  
Db 726 TEGRYFESSDLTWSLWPIGVKFRFSDNDIATSVNLAVAPDLVRSNPDCKTSLVSPPT 785  
Qy 959 --WTTDAFLARHGVVVVGRGSMYASLTNIEVYGHGRYEDRAGRGYGLSAGSVRF 1012  
Db 786 AVMLTKATNLARHAFIVKAGNYLSLSSNPFISQFGFELGSSRTYNVDLGSKIQF 841  
RESULT 20  
PMP1 CHLPN STANDARD; PRT; 922 AA.  
AC Q929G5; Q9KLY9; Q924H9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

10-MAY-2005 (Rel. 47, Last annotation update)  
probable outer membrane protein pmp1 precursor (Polymorphic membrane protein 1) (Outer membrane protein 6).  
Name=pmp1; Synonyms=omp6; OrderedLocusNames=CPn0005, CP0770, CpB0006;  
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.  
OX NCBI\_TaxID=83558;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CWL029 / VR1310;  
RX MEDLINE=20007584; PubMed=10539856;  
RA Christensen G., Boesen T., Hjerno K., Dagaard L., Mygind P., Madsen A.S., Knudsen K., Falk E., Birkelund S.;  
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity";  
RL Am. Heart J. 138:8491-8495(1999).  
RN (2)  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=CWL029;  
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;  
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W., Olinger L., Grimwood J., Davis R.W., Stephens R.;  
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."; Nat. Genet. 21:385-389(1999).  
RN (3)  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=AR39;  
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J., Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C., Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F., McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39."; Nucleic Acids Res. 28:1397-1406(2000).  
RN (4)  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=J138;  
RX MEDLINE=20330349; PubMed=10671362; DOI=10.1093/nar/28.12.2311;  
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishi K., Hattori M., Kuhara S., Nakazawa T.;  
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA."; Nucleic Acids Res. 28:2311-2314(2000).  
RN (5)  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=TW-183;  
RX Geng M.M., Schumacher A., Muehlbauer I., Bensch K.W., Schaefer K.P., Schneider S., Pohl T., Essig A., Marre R., Melchers K.;  
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis."; Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).  
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC EMBL; AJ133035; CAB37075.1; -; Genomic DNA.  
CC EMBL; AB001585; AAD18163.1; -; Genomic DNA.  
CC EMBL; AB002237; AAF38570.1; -; Genomic DNA.  
CC EMBL; BA000008; BAA98215.1; -; Genomic DNA.  
CC EMBL; AB017157; AAP97939.1; -; Genomic DNA.  
CC FIR; B72131; B72131.  
CC FIR; B86491; B86491.  
CC TIGR; CP0770; -.  
CC InterPro; IPR005546; Auto transpbeta.  
CC InterPro; IPR011427; ChlamPMP\_M.

DR	InterPro; IPR003368; Chlamydia_PMP.
DR	Pfam; PF03797; Autotransporter; 1.
DR	Pfam; PF02415; Chlam_PMP; 7.
DR	Pfam; PF07548; ChlamPMP_M; 1.
DR	TIGRFams; TIGR01376; POMP_repeat; 5.
KM	Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT	SIGNAL 1 26 Potential.
FT	CHAIN 27 922 Probable outer membrane protein pmp1.
FT	CONFLICT 14 14 F -> L (in Ref. 1).
FT	CONFLICT 251 251 A -> P (in Ref. 5).
FT	CONFLICT 375 375 Y -> C (in Ref. 1).
FT	CONFLICT 606 606 D -> N (in Ref. 1).
FT	CONFLICT 836 836 S -> P (in Ref. 5).
SQ	SEQUENCE 922 AA; 100458 MW; DFF2AB633AB031C CRC64;
Query Match 18.4%; Score 967; DB 1; Length 922;	
Best Local Similarity 28.0%; Pred. No. 1.6e-48;	
Matches 295; Conservative 144; Mismatches 380; Indels 236; Gaps 32;	
Qy	31 MYPQGIYDGETLTVPFYTVIGDSPGTTVFAGBELTKNLNDNSAAALPLSCFGNLLGSFT 90
Db	31 LTPEDSPHGDSQNAERSINV---QAGDVYSLTGDVSIENVDS--ALNKACFNVTSGSVT 85
Qy	91 VLGRGHSLTPTENIKTST--NGAALSNSAADGLFTIEGPKELSPNCNSLLAVLPAATTNK 148
Db	86 FAGNHGGLYFNNISSGTYKEGAVLCCDDPQATARFSGFSTLSF----- 128
Qy	149 GSOTPTTSTPSNGTIYSKTDLLILNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVF 208
Db	129 -IGSP--GDIKEQCLYSKNALMLNLYVVRPEQNSKTKGAGISGANVTIVGNVDSVSP 185
Qy	209 QENTAQADGGACQVVTSPSAMENAPITAFVANVAGVGGGIAAVQDGOQGVSSSTSTEDP 268
Db	186 YQNAATFGG-----ALHSGSPQLQIAVNOAER----- 212
Qy	269 VVPSRNTAVFDCGNVARVGGIYSYGNVAFNNKTLFLNNVASPVYIAAKQPTSGQAS 328
Db	213 ---FAQNTAK--NGS---GGALYSDDGIDIDQNAVYLFRENEALTTAI----- 252
Qy	329 NTSNNYDGGGAI FCKNGAQAGSNNGS-----VSFDGEGVVFPSSNVAAKGGAIVYAKK 382
Db	253 -----GKGGAVCC-----LPTSGSGTPVPIVTFSDNKKQLVFNHNSIMGGAIVYAKK 299
Qy	383 LSVANCPVQVPLRNIA-----NDGGATYLGESGELSLSDYDGIIFDGNLKTAKENAAD 437
Db	300 LSISSGGPTLFINNISVANSQNLGGAITDGGHLSLAEKGTITFGN-----RTSLPP 354
Qy	438 VNGVTVSSQAISMGGGKITTLRKAGHQILFNDPIEMANGNNOQAGSKLLKIN---DG 494
Db	355 LNGIHLLQNA-----KFLKQARNGYSIEFYDPI-----TSEADGSTQLINGDPPKN 401
Qy	495 EGYVTGDIIVA-----NGSSTLVQNTVIEGRIVLREKAKLSVNSLSQTSGLS-LYM 543
Db	402 KEYTGITLPSGEKSLANDPRDFKSTIPQNVLSAGYLVIKEAGVATVSKFTQSPGSHLV 461
Qy	544 EAGSTWDFVTPQPQPAPANOLLITLSLHLSLSLLANNAVTPNTPPQAQDSHPAIVG 603
Db	462 DLGT-----KLASKEDAITGLAIDISL-----SSSSTAIVIK 496
Qy	604 STTAG-----SVTISGPIF-----FEDLDDTAYDRYDMLGNSQKINVLKQLGTGTPKANAP 654
Db	497 ANTANKQISVTDSTIELISPTGNAYEDLR-----WRNSQTPPLLSEL---PGAGGS 543
Qy	655 SDITLGNEM---PKYGYQGSWKLAWDPNTANGPYTLKATWTKGYNGPDPERVASLVPNS 711
Db	544 VTTVAGDPLVSPHYGFGQGNWKLAWTGTGNKVGEF----FWDKINIKYKPRPEKGNLVPI 599
Qy	712 LWGSLIDIRSAHSIAQASVDGRS*CRGLWTVSGVSNFFVHDDRDLAQGVRYISGGVSYLGN 771
Db	600 LWNVANVDRSLMQVQETHASLSQTRDGLNLDGICGNFFHVSASEDNIRYHNSGGTVLSVN 659
Qy	772 SYFGSSMF-GLAFTEVPGRSKDYVYVCRNSHHACTIGSVYLS*QQALCCGYLFI-----GDA 824

1004 LSAGSRVRF 1012  
839 IDLGSKIQF 847

RESULT 22  
P77792\_CHLAB PRELIMINARY; PRT; 839 AA.

AC P77792\_Q5LS95;  
01-FEB-1997 (TrEMBLrel. 02, Created)  
01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE POMP90B precursor (POMP90A precursor) (Polymorphic outer membrane protein).  
DN Name=pmp17G; Synonyms=pomp90B; OrderedLocusNames=CAB598;  
OS Chlamydomophila abortus.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiaophila.  
NCBI\_TaxID=83555;  
[1]  
RN NUCLEOTIDE SEQUENCE.  
RP  
RC STRAIN=S26/3;  
RC MEDLINE=98187897; PubMed=9529048;  
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;  
RT "Molecular cloning and characterization of the genes coding for the  
RT highly immunogenic cluster of 90-kilodalton envelope proteins from the  
RT Chlamydia psittaci subtype that causes abortion in sheep."  
RL Infect. Immun. 66:1317-1324(1998).  
[2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=S26/3;  
RC MEDLINE=96406378; PubMed=8810511; DOI=10.1016/0378-1097(96)00281-9;  
RA Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;  
RT "Identification of a multigene family coding for the 90 kDa proteins  
RT of the ovine abortion subtype of Chlamydia psittaci."  
RL FEMS Microbiol. Lett. 142:277-281(1996).  
[3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=S26/3;  
RC PubMed=15937807; DOI=10.1101/gr.3684805;  
RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,  
RA Livingstone M., Cerdeno-Tarraga A.-M., Harris B., Doggett J.,  
RA Ormond D., Mungall K., Clarke K., Fellwell T., Hance Z., Sanders M.,  
RA Quail M.A., Price C., Barrall B.G., Parkhill J., Longbottom D.;  
RT "The Chlamydomophila abortus genome sequence reveals an array of  
RT variable proteins that contribute to interspecies variation."  
RL Genome Res. 15:629-640(2005).  
DR EMBL; U65943; AAC15924.1; -; Genomic DNA.  
DR EMBL; U65942; AAC15922.1; -; Genomic DNA.  
DR EMBL; CR848038; CAH64045.1; -; Genomic DNA.  
DR GO; GO:0019867; C:outer membrane; IEA.  
DR InterPro; IPR005546; Auto transporter.  
DR InterPro; IPR006315; Auto transporter.  
DR InterPro; IPR011427; ChlamPMP\_M.  
DR InterPro; IPR003368; Chlamydia\_PMP.  
DR Pfam; PF03797; Autotransporter; 1.  
DR Pfam; PF02415; Chlam PMP; 4.  
DR Pfam; PF07548; ChlamPMP\_M; 1.  
DR TIGRFAMs; TIGR01414; autotrans bar1; 1.  
DR TIGRFAMs; TIGR01376; POMP\_repeat; 3.  
KW Complete proteome; Signal.  
FT SIGNAL 1 16 Potential.  
FT CHAIN 17 839 POMP90B.  
SQ SEQUENCE 839 AA; 89825 MW; 4581C7CBAP7FP4C4 CRC64;

Query Match 18.0%; Score 949.5; DB 2; Length 839;  
Best Local Similarity 27.3%; Pred. No. 1.5e-47;  
Matches 291; Conservative 153; Mismatches 341; Indels 281; Gaps 36

QY 1 MQTSHFKPLSMILAYSCCSLGGVAAEIMVPGQIYDGETLTVFPFYTVIGDPSGTTVF 60  
DB 1 MKHPYVFWLISSL-FANSLSFANDAQTALFSDSYNGVNTSEFP-QVKETSSGTTVT 57

QY 61 SAGELTLKNLSDNSAALPLSCFNLGSLFTVLGRHSLTFENIRTSITNGAALSNAAAGD- 119









```

Db      358 KSFQKQKSLFWHFGT-----RLQTNL-SIKNLHLDLAKIAA-TAV----- 399
QY      591 NPQAQSHPAVIGSTTAGSVTSGTFFEDLDDTAYDRYDVMGSGNOKINVLKQLGTQKPP 650
Db      400 -----EIAATADAAAEICGPMVWH-VDEIFPYNQALADSLSPFCLHVR----- 443
QY      651 ANAP--SDLTGNEMP-----KYGVQGSWKLAWDP-----NTANNPYTLKATW 692
Db      444 --APHLDNITV-DDVPLIPITTMETHRGYQGWTSWEEHEHMIFGNVSTQPNKQKSLVW 500
QY      693 TKTGYNP-----GPERVASLVPNSLWGSILDIIRSAHSAIQASVDGRSYCKRLWVGVSVP 748
Db      501 NPSGYIPFVGCGEFTTSLVPSNLWMLFLDTFAQAQAI--TNAQSPGNGIWLISLTSNF 558
QY      749 YHRRDALGGYRYISGYSGLAN--SYFGSSMFLGAPTEVFGRSKDYVCRSHHACIGSV 807
Db      559 RKGSTENNHGFRHKSBSGYVAGGKFTQLDDIFSVGICQLFGRSKDFGSAKSKAPFSGSL 618
QY      808 YLSTQOAL--CGSYLFG-----DAPIRASVGFGRHQMKTSTYTPAE 845
Db      619 YAHHSRYLLPITRFLAGTSGRSQRFSLRIPKDPPIINFDAI--SYSGRNHMKVYADHS 676
QY      846 ESDVRDNNCLAGEICAGLPVITPSKLYLNLRLPFVQAEFSYADHESFTEBGDQARAFK 905
Db      677 QTTSWNTGYGAQIGSSLPCLDVSHTFPQVSPFVKLHWIYAHQVQEQEGIKERSEN 736
QY      906 SGHLLMSVPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETTLIS-HQETWTTDAP 964
Db      737 NSNLKSLPIGLKI-QGSLHLHLSYELGTWYIADLYRCNPESVTSLSGGLPWTTTAA 795
QY      965 HLAHGVVVRGSMYASLTNIEVYGHGRYEDRDSRGYGLSAGSRVP 1012
Db      796 NLGKQALLQGSGLNLTSHINIPAQGTVEFRSSYSYAMDAGSRVHP 843

RESULT 26
PMPI_CHLMU
ID PMPI_CHLMU STANDARD; PRT; 867 AA.
AC Q9PLA1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable outer membrane protein pmp1 precursor (Polymorphic membrane protein I).
DE Name=pmp1; OrderedLocusNames=TC0267;
DE Chlamydia muridarum.
DE Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MOPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (potential).
CC -1- SIMILARITY: Belongs to the pmp outer membrane protein family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE002294; AAF39136.1; -, Genomic_DNA.
DR PIR; F81721; F81721.

```

```

DR TIGR; TC0267; --
DR InterPro; IPR005546; Auto transpbeta.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam PMP; 6.
DR Pfam; PF07548; ChlamPMP M; 1.
DR TIGRPFAMs; TIGR01376; POMP repeat; 6.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 29 Potential.
FT CHAIN 30 867 Probable outer membrane protein pmp1.
SQ SEQUENCE 867 AA; 95018 MW; 557994185A9B5652 CRC64;

Query Match 15.0%; Score 789; DB 1; Length 867;
Best Local Similarity 26.1%; Pred. No. 5.2e-38;
Matches 268; Conservative 148; Mismatches 411; Indels 200; Gaps 31;

QY 18 CCSLNGGGAARIMVPOGIYDETLTVTFPVTYVIGDPSTGTVFSAGELTKLNDLSIAL 77
Db 9 CC-----LCATILSPTAILFGQD-ALDKSALITKNPNSIVCTFLEDCTMENFSPALLSH 61

QY 78 -----PLSCFNGLLGSFTVLGRGHSILTPENIRTSNGAALSNSAADGLFTIEGFKELSPS 132
Db 62 ARQDDPLVIYIGNT-----HWFVSNLHPSTNHEERFLKEKGD--LSIQDFRLST 109

QY 133 NCNLSLAVLPAATTNKGSTPTTTGTPNSGTIYSKT-DLLLLNNEKFSYSLVSDGGA 191
Db 110 DCSS-----STEDSPS--ILYHKNQQLFLRNGNWSFYRNHSESGGA 150

QY 192 IDAKSLTVQGISKLQVQENTAQADGGACQVVTSPFSAMANEAPIAFVANAVGVRGGIAA 251
Db 151 LSTDALFLQHNYLFTNFENSSAKNGAIOAQT----- 183

QY 252 VDDGQGVSSSTSTEDPVVPSRN-TAVEFDGNVAVRGGIYSYGNVAFLNNGKTLPLNN 310
Db 184 -----LSLSRNVSSLSFSRNRANLNGAICCCQNLISGNVNPILFF-- 223

QY 311 VASPVYIAAKQPTSGQASNTSNYDGGAIKCKNGAQAAGSNNSGVSFDEGVVVFSSNV 370
Db 224 -----TNSALNGGAIKCCINEQLSEKGLSLAYNQE--TLFSGNS 262

QY 371 AAKGGAIYAKKLSVANCGPVQFLRNANDGGAIVYGESGELSLSADYDIIPDGNLKR 430
Db 263 AKEKGAIVTKHMLRHNGPVSFVNSAKLGAIAIQSGGSLIIAGGSSVLPQNNSCHP 322

QY 431 AKENAAVNGVTVSSQAISMGGKTTIRAKAGHOILFNDPIEMANGNNQPAQSKLLK 490
Db 323 SDQGTV-----RNAIYLEKNALLSSLEARHG-DILFFDPVIVQEVSPFETTSALT 373

QY 491 INDGEGYTGDIVP-----ANGSSTLYQNVITBQGRIVLRKAKLSVNSISQTG 538
Db 374 LRIQTWTRAVIFSSNLSKEKTEANLISKQOPIELQSGCLVLDKRVILSAPLSQAP 433

QY 539 GS-LYMEAGSTWDFVTPPQPPQPPAAANQITISNLHLSLSSLLANNNAVNTPTNPPA 597
Db 434 QALLVMDVGT-----LTTSDKLTLTSLPLHSI----- 463

QY 598 HPAVIGSTTAGSVTISGP-----IPFEDLD-DTAYDRYDVLGSGNQK-INVLKQLGT 649
Db 464 -----DTENSVSIOQPTLSIQKIFLSNSENHYENVLLSKQKQDIPLLSLPGLPH 516

QY 650 PANAPSLDTLGNEMPKYGVGSKWKLAWDPNTANNPFTLKATWTKTGYNPGPERSVSL 709
Db 517 P-----DLPDGNLSHFQYQGDWNFSWQTSDORE----TLVANVTANSYIHPERQ 568

QY 710 NSLWGSILDIRSAHSAIQASVDGRSYCKRLWVGVSVPFFY-HDRDALG-QGYRYISG 767
Db 569 NTLWNTYSMDQAVQSMINTTAQCGAYLFGTWGSAVSNLFFSHGNSGKSTDNWKRSL 628

QY 768 LGANSY-FGSSMFLGAPTEVFGRSKDYVCRSHHACIGSVYSTQALCGSVLFGDAFI 826
Db 629 FGISTHSLDDHSFCLAAAGQLFGKSSDSFVTSADTTSYIAI-----QTATSLIK 682

```

```

Qy 827 RASVGFQGHMKTYS-TFAESDVRWNNCLAGEIGAGLPIVITPSKLYLNELRPFVQAE 885
Db 693 QACYNESHEIKTKRFSKSGFGAMHSAVSGEIGASIPVNSGGLF-SFSIFSFKIQ 741
Qy 886 FSYADHESFTREGDQARAFKSGHLNLNLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAVRTI 945
Db 742 GFSGKQDGFESRGEARAFADSSFTNISLPVGAIEPKKSQKTRNYHFLGAYIQDLKRCV 801
Qy 946 SGTETILLSHGETWTTFDAPLHAGVGVVRGSMYASLTNSNIEVYGHGRYEYDASRGYGLS 1005
Db 802 ESGPVTLTKNSVTWDPANLDSRAMFRLTNQAL-HRFQTLVNMSTMRLGQSYSLD 860
Qy 1006 AGSRVRF 1012
Db 861 LGTYRPF 867

RESULT 27
ID Q8VU49_CHLPS PRELIMINARY; PRT; 602 AA.
AC Q8VU49;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative polymorphic membrane protein (fragment).
OS Chlamydia psittaci (Chlamydiales; Chlamydiaceae; Chlamydia).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LLG;
RA Laroucau K., Souriau A., Rodolakis A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243418; AAL36962.1; -; Genomic DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Auto transporter.
DR InterPro; IPR005546; Auto transporter.
DR InterPro; IPR011427; Chlam_PMP_M.
DR Pfam; PF03797; Auto transporter; 1.
DR Pfam; PF07548; Chlam_PMP; 1.
DR TIGRFAMs; TIGR01414; Autotrans_barl; 1.
FT NON TER
SQ SEQUENCE 602 AA; 65561 MW; CA486CFACEC131B2 CRC64;

Query Match 15.0%; Score 787.5; DB 2; Length 602;
Best Local Similarity 33.0%; Pred. No. 3.8e-38;
Matches 210; Conservative 99; Mismatches 250; Indels 77; Gaps 21;

Qy 419 GDIIIPDGNLKTAKENADVNGVTSSQAIISWGGKITTILRAKAGHQILFNDPIEMANG 478
Db 2 GDITPDGNKIITTSRSSS-----TVKRNISISLGGKPTKLNAKEGFGIFPYDPIANTGD 56
Qy 479 NNQPAQSSKLKINDGEG-----YTGDIVFA-----NGSSTLYQNVTIEQGR 521
Db 57 TNTS-----IELNAKEGGSTTYTGKIVFSEKLSDEKVDNLKSYFTQPLKIGAGSL 110
Qy 522 VLREKAKLSVNSLSQSGS-LYMEAGSTWDFVTFPQPPQPPAANQLTSLNLHLSLSLL 580
Db 111 VLKDGVTLEAKKVGQTDGSTVMDLGT-----LQTSSTSGETITLTNLNDINVAS-L 160
Qy 581 ANNAVTPNPPADQSHPAVIGSTAGSVTISGPIPEDLDDTAYDRYDHLGNSOKINV 640
Db 161 GGGGVADPFAKVAQAS-----GKT-----VTINA-VNLVDVDGNAYE-YPLATSQPPTA 209
Qy 641 LKLQGTGKPPANAPSDTLTGNEMP--KYGYGSGKLAWDPNANNGPYTLKATWTKTGYN 698
Db 210 IIAKAGSGTTTTTD-NLKNYTPHYGQGNVTYVWKLGTSAQ-BETATLTWEQTDYS 267
Qy 699 GGPBRVASLVNSLWGSILDIRSAHSAIQASVDGRSYRCGLVWVGNSNFFYHDDRDLGQG 758
Db 268 FNPBRQGLVPNTLWGSFSDIRAIQNLIDISVNGADYRRGFWVSGLGNFLHKSGSNTRK 327
Qy 759 YRISGGVSLGANSYFGS-SNFGLAFTVFGRSKDYVYVCRSNHHACIGSVYL-----ST 811

```

```

Db 328 FRHSAGYALGVTAQTSTEDVFSAAFCQLFKOKDKDYFVSKNSSNIYAGSIYQHISYNA 387
Qy 812 QQALCGSYLFGDAFI-----RASVGFQGHMKTYSY-----PABESDVRWNNCLAGE 859
Db 388 WONLQSTIGAEAPLVLAQLTYCHASNKMTNTVTPKNTVLTSEIKGDMGNDGCGVE 447
Qy 860 IGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTS-EGDQARAFKSGHLNLNLSVPVGV 918
Db 448 FGAMAPIE-TPSSFLDFRYSFPFLQLQLVHAHQDDFKENSDQGRYPFESSNLTNLSMPTGI 506
Qy 919 KFDRCSSSTHPNKYSFMAAYICDAVRTISGTETILL--SHQETTTTDFLHARHGCVVVRGS 976
Db 507 KPERFAYNDVASVHLTAAYAPDIVRNPDCASLVLSVPTSAVWVTKANLARSAPMLQAG 566
Qy 977 MYASLTNSNIEVYGHGRYEYDASRGYGLSAGSRVRF 1012
Db 567 NYLALSHNMLFSGQFGEIRGSSRTYVNDLGSKIQF 602

RESULT 28
ID Q823X3_CHLCV PRELIMINARY; PRT; 843 AA.
AC Q823X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family protein/autotransporter, putative.
OS OrderedLocusNames=CCA00280;
OC Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T., Heidelberg J.F., Holtzaple E.K., Khouri H.M., Federova N.B., Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J., White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavoil P.M., Fraser C.M.;
RA "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016995; AAP05031.1; -; Genomic DNA.
DR TIGR; CCA00280; -;
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto transporter.
DR InterPro; IPR011427; Chlam_PMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR006626; Pbh1.
DR Pfam; PF03797; Auto transporter; 1.
DR Pfam; PF07548; Chlam_PMP; 1.
DR Pfam; PF02415; Chlam_PMP; 4.
DR SMART; SM00710; Pbh1; 2.
DR TIGRFAMs; TIGR01376; POMP_repeat; 3.
KW Complete proteome.
SQ SEQUENCE 843 AA; 90476 MW; 0BB240F4687AFC7A CRC64;

Query Match 14.9%; Score 785.5; DB 2; Length 843;
Best Local Similarity 26.1%; Pred. No. 8.1e-38;
Matches 270; Conservative 131; Mismatches 351; Indels 281; Gaps 37;

Qy 37 YDGETLTVSPYTVIGDPSTGTVFVSAGELTKMLNLSIALPLSCFNLGSGFTVLGRGH 96
Db 35 FDSGTAGQFTSKQSTDAGGTTTLTGDDVTIIVKTTSPA-NTSCFNSGNLFTTGANH 93
Qy 97 SLTFENIRTSNAGLSNSAADGLFTIEGPKELSFNSCNLSLLAVLPAAATNKGSTPTTT 156
Db 94 SLIFEDIISTAQGAAS-----ANT----- 113

```



```

Db 530 LWNFLPLDTRFQQAIEK--HAVSSGNGIWSMTNTPSQGSSNNHNGFRHKSSGGYTAGGK 587
Qy 772 -SYFGSSNFWGLAPTEVFGRSKDYVCRNHNHACIGSVYLSQQALCG--SYLFGDAPTR- 827
Db 588 IQTLQDDIFSVFSQPKRSKDFGSATSKDTFLSGSIYAQHSRLLPTMRFLAGTSTYRP 647
Qy 828 -----ASYGFGNOHMKTSYTFABESDVRDNNCLAGEICAGLPVITP 870
Db 648 RLLISIPKNLPIINFVLVSYSDSNHMKVKNSNRTQTRGSSNTPGYSAQIGSSIPFVLVDV 707
Qy 871 SKLYNELRPFVQAFESYADHESFTBEGDQARAFKSGHLLNLSVPVGVKFPDRCSSTHFNK 930
Db 708 SHTFPQYVSPFKLHWIAHQVQFOEQGKRSFNNSNLKLSLPLGLKI-QGQTLRLS 766
Qy 931 YSFMAAYICDAYRTISGTETILLS-HQETWTTDAPHLARHGVVVRGSMYASLTNIEVYG 989
Db 767 YEFTGMWITDLYRCNPEGSVTSLSIGLLPWTITGNTLSKQAALLQGSNSISLTSHINIFA 826
Qy 990 HGRVEYRDASRGYGLSAGSRVRF 1012
Db 827 QGTVEFRSSYSYNDMDFGSRVHF 849

RESULT 30
Q4VWR2_CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR2
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TW-3, TW-448, TW-5, and Apache-2;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299430; AAQ74446.1; -; Genomic DNA.
DR EMBL; AY299442; AAQ74458.1; -; Genomic DNA.
DR EMBL; AY299448; AAQ74444.1; -; Genomic DNA.
DR EMBL; AY299449; AAQ74445.1; -; Genomic DNA.
DR InterPro; IPR005546; Auto.transp.bet.
DR InterPro; IPR011427; ChlamPMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
SQ SEQUENCE 878 AA; 95489 MW; 927611D9B0B636F7 CRC64;

```

```

Query Match 14.9%; Score 784.5; DB 2; Length 878;
Best Local Similarity 26.5%; Pred. No. 9.8e-38;
Matches 267; Conservative 146; Mismatches 392; Indels 203; Gaps 33;

Qy 53 DPSTGVTFPSAGELTKLNDLSIAALPLSCFNGLL-----GSFTVLGRGSLTFE 101
Db 26 DPLGETA-----LITKPNHVVCFFEDCTMESLFPALCAHASQDDPLVVLGNSYCWFS 80

Qy 102 NIRTSTGAALNSAAAGLFTIEGKELSPNCNKLAVLPAATNKGSPPTTTPSN 161
Db 81 KLHITDPKEALFKBKGD--LSIQNFRFLSPFTDCSS-----KESSPS- 119

Qy 162 GTIYSKT-DLLLNNEKFSFYNLSVSGDGGDAIDAKSLTVQGISKLCVFQENTQAQDGAC 220
Db 120 -IIHKQKQLSLRNGNSFCRNHABGGGAISADAFSLQHNLYLFTAFENSSKNGGAI 178

Qy 221 QVWTSFSAWANEAPAFVANVAVRGGGIAAQQDQGGQVSSSTSTEDPVVFSRNTAVEF 280
Db 179 QAQT-FSLSRNVSPISFARNADLNGGAICC----- 208

```

```

Qy 281 DGNVARVGGGIYSYCNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNTVDGGA 340
Db 209 -----SNLCSGNV-----NPLFF-----TGNSATNGCAI 233
Qy 341 FCKNGAAGSNNSGSVFDGEGVVFFSNVAAGKGAIAKAKLSVANCGPVQPLRNIAND 400
Db 234 CCI--SDLNTSEKSLACNQETLFASSNAKEKGAIAKHMVLRVNGPVSPINNSAKI 291
Qy 401 GGAIYLGESGELSADYGDIIFDGNLKRKTAKENADVNGVTVSSQAISMGSQGGKITTIR 460
Db 292 GGAIYLGESGELSILAGEGSLFQNNQRTSDQGLVR-NAIYLEKDAI-----LSLE 343
Qy 461 AKAGHQILFNDPI-EMANGNNQPAQSSKLLKINDGEGYT-----DIVPA----- 504
Db 344 ARNG-DILFFDPIVQESSKESPLSSLOASVTSPTPATASPLVIQTSANRVSIFSGERL 402
Qy 505 -----NGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
Db 403 SEBEKTPDNLTSQLPIELKSGRLVKDRAVLSAPLSQDPQALLIMEAGTS----- 455
Qy 556 PPOQPPAANQLITLSNLHLSLSLLANNAVTPNPTPPAODSHPAVIGSTTAGSVTISGP 615
Db 456 -----LKTSSDLKATLSIPLHSL-----DTEKSVTIHAP 485
Qy 616 -----IPFEDL-DDTAYDRYDMLGNSQKINVLKQLGTQPPANAPSDLTIGNEMPKYGY 668
Db 486 NLSIQKIFLNSGDNENFENVLLSKEQN-----NIPLLTLPKEQSHLHPDGNLSSHFGY 541
Qy 669 QGSMKLAWDPTANNNGPYTLKATWTKGYNPGPERVASLVPNSLWGSILDIRSAHSAIOA 728
Db 542 QGDWTFSW--KDSDEG-HSLIANWTPKNYVPHPERQSTLVANTLWNTYSMDQAVQSMINT 598
Qy 729 SVDGSRYSRGLWVGVSNGFFY-HDRDALG-QGYRIISGGYSLGANSY-FGSSMFGLAFT 785
Db 599 TAHGAYLFGTWGSAVSNLFAHDSGKPIDNWHHRSLSGLYLFGLSTHSLDHSFCLAAGQ 658
Qy 786 VFGSKDYVYVCRNHNHACIGSVYLSQQALCGSYLFGDAPFIRASYFGNGHMTSY-TPA 844
Db 659 LLGKSSDSFITSTE-----TTSVIATVQAQLATSLM-KISAQACYNESIHELKTKYRSFS 712
Qy 845 EESDVRDNNCLAGEICAGLPVITPCKLYNELRPFVQAFESYADHESFTBEGDQARAP 904
Db 713 KEGFGSMHVSVAVSGEVCALPIVSGSGLF-SSFSIFSKLQSGFGTQDGFEESSGEIRSF 771
Qy 905 KSGHLLNLSVPVGVKFPDRCSSTHFNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAF 964
Db 772 SASSFRNISLPIGITPEKSKQTRNYYPFLGAYIQDLKRDVESGPPVLLKNAVSMADPMA 831
Qy 965 HLAHGVVVRGSMYASLTNIEVYGHGRYRYDRASRGYGLSAGSRVRF 1012
Db 832 NLASRAYMFLTNQAL-HRLQTLNLSVYVLRQGSYSHSLDLGATHRP 878

RESULT 31
Q4VWR2_CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR2
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Bour;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299443; AAQ74459.1; -; Genomic DNA.

```

```
SQ SEQUENCE 878 AA; 95629 MW; CA385BFC8BAC1C17 CRC64;
Query Match 14.8%; Score 781.5; DB 2; Length 878;
Best Local Similarity 26.6%; Pred. No. 1.5e-37;
Matches 269; Conservative 146; Mismatches 388; Indels 207; Gaps 35;

QY 53 DPSGTTVFSAGELTLKNDLSAALPLSCFGNLL-----GSFTVLGRGHSLTPE 101
DB 26 DPLGETA-----LLTKNPNHVCTTFEDCTMESLFPALCAHASQDDPLYVLGNSCWVFS 80

QY 102 NIRTSTNGAALSNSAADGLFTIEGPKELSFNSCNLSLLAVLPAATTNKGSTPTTTSTPSN 161
DB 81 KLHITDPKEALPKEGD--LSIQNFRFLSFTDCSS-----KSSSPS- 119

QY 162 GTIYSKT-DLLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOENTAOADGGAC 220
DB 120 -IIHQKNGQLFLRNNGSMFSCRHAEGSGGAIADAFSLQHNYLFTAPEENSKNGGAI 178

QY 221 QVVTSPSAMANEAPAFVANVAGVGGGIAAQQDQGGVSSSTSTEDPVVPSRNTAVEF 280
DB 179 QAQT-FSLSRNVSPISFARNRADLNGAIC----- 208

QY 281 DGNVARVGGGIYSGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNNGDGGAI 340
DB 209 -----SNLICSGNV-----NPLFF-----TGN SATNGGAI 233

QY 341 FCKNGAQSNNSGSVSFDGEGVFPFSSNVAAGKGAIIYAKKLSVANCGPVOFLRNAND 400
DB 234 CCI--SDLNTSEKGSLSLACNQETLFPASNSAKEKGAIYAKHMLRYNGPVSFINNSAKI 291

QY 401 GGAIVYGESEGLSLADYGDIIIDGNLKTAKENADVNGVTVSSQAIISWGSGGKITTLR 460
DB 292 GGAIAIQSGGSLILAGEGVLFPQNSQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343

QY 461 AKAGHOILFNDPT-EMANGNNOQAOSKLLKINDGEGYTG-----DIVFA----- 504
DB 344 ARNG-DILFPDPIVQSSSKESPLPSLSQASVTSPTPATAPLVIQTSANRSVIFSSERL 402

QY 505 -----NGSSTLYQNVTIEQRIYVLRKAKLSVNSLSQTGGS-LYMBAGSTWDPVTPQ 555
DB 403 SEEEKTPDNLTSQLOQPIELKSGRLVKDRAVLSPSLSDQDQALLIMEAGTS----- 455

QY 556 PPOQPPAANQLITLSNLHLISLSLLANNVTPPTNPPAQDHPAVIGSTTAGSVIISGP 615
DB 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP 485

QY 616 -----IFPEDL-DDTAYDRYDMLGSKNKLNLVLKQLGTGKPPANAPSDLTL--GNEMPKY 666
DB 486 NLSIQIKIFLNSGDNFENVELLSKEQN-NIPLLT-----SKEQSHLHLPDGNLSSH 539

QY 667 GYQGSWKLAWDPNTANGPYTLKATWTCTGYNPGERVASLVPNSLWGSITLDRSAHSAI 726
DB 540 GYQGDWTFPW--KDSDEG-HSLIANWTPKNYVPHPERQSTLVANTLWNTYSDMQAVQSMI 596

QY 727 QASVGRSVCRLVWGSVSNFFY-HRDALG-QGYRIYISGYSGLANSY-FGSSMFGCLAF 793
DB 597 NTTAHGGAIVLFGTWGSAVNSLFYAHDSGKPIDNWHRSIGLYLFGISTHSLDHDHSCFLAA 656

QY 784 TEVFGSKDVVCRSNHHACIGSVLSTQALCGSVLFGDAFIRASVFGPGNQMKTSY-T 842
DB 657 GQLLGKSSDSFITS-ETTSYIATVQAQLATSLM-KISAQACYNESIHELKTKYRS 710

QY 843 FASESVRWNDCIAGIEAGLPIVTPPSKLYLNELPFPVQABFSYADHESFTEEGDQAR 902
DB 711 PSKEGFGSWHSAVSGEVCASIPVNSGGLP-SSPSIFSKLQFGSQDQGFESSSGEIR 769

QY 903 AFKSGHLNLT.SVPVGVKFDRCSTHPNKNYSFMAIYICDARTISGTETTLLSHQETWTD 962
DB 770 SFSASSPRNLSLPIGTIFTEKKSQKTRNYIYFLGAYIQDLKRDVKSQGPVTLTKNAVSADAL 829

QY 963 AFHLARHGVVVRGSMYASLTSNIENVYGHGRYEYRDASRGYGLSAGSRVF 1012
DB 830 MANLDSRAYMFLRLNQAL-HRLQTLNLVSVLGRQSHSYSLDLGTTTRF 878
```

## RESULT 32

```
Q4VWS3_CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWS3;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IC-Cal-3;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299433; AAQ744448.1.; Genomic DNA.
SQ SEQUENCE 878 AA; 95569 MW; 30907BB98BADF659 CRC64;
Query Match 14.8%; Score 779.5; DB 2; Length 878;
Best Local Similarity 26.6%; Pred. No. 1.9e-37;
Matches 269; Conservative 146; Mismatches 388; Indels 207; Gaps 35;

QY 53 DPSGTTVFSAGELTLKNDLSAALPLSCFGNLL-----GSFTVLGRGHSLTPE 101
DB 26 DPLGETA-----LLTKNPNHVCTTFEDCTMESLFPALCAHASQDDPLYVLGNSCWVFS 80

QY 102 NIRTSTNGAALSNSAADGLFTIEGPKELSFNSCNLSLLAVLPAATTNKGSTPTTTSTPSN 161
DB 81 KLHITDPKEALPKEGD--LSIQNFRFLSFTDCSS-----KSSSPS- 119

QY 162 GTIYSKT-DLLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOENTAOADGGAC 220
DB 120 -IIHQKNGQLSLRNNGSMFSCRHAEGSGGAIADAFSLQHNYLFTAPEENSKNGGAI 178

QY 221 QVVTSPSAMANEAPAFVANVAGVGGGIAAQQDQGGVSSSTSTEDPVVPSRNTAVEF 280
DB 179 QAQT-FSLSRNVSPISFARNRADLNGAIC----- 208

QY 281 DGNVARVGGGIYSGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNNGDGGAI 340
DB 209 -----SNLICSGNV-----NPLFF-----TGN SATNGGAI 233

QY 341 FCKNGAQSNNSGSVSFDGEGVFPFSSNVAAGKGAIIYAKKLSVANCGPVOFLRNAND 400
DB 234 CCI--SDLNTSEKGSLSLACNQETLFPASNSAKEKGAIYAKHMLRYNGPVSFINNSAKI 291

QY 401 GGAIVYGESEGLSLADYGDIIIDGNLKTAKENADVNGVTVSSQAIISWGSGGKITTLR 460
DB 292 GGAIAIQSGGSLILAGEGVLFPQNSQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343

QY 461 AKAGHOILFNDPT-EMANGNNOQAOSKLLKINDGEGYTG-----DIVFA----- 504
DB 344 ARNG-DILFPDPIVQSSSKESPLPSLSQASVTSPTPATAPLVIQTSANRSVIFSSERL 402

QY 505 -----NGSSTLYQNVTIEQRIYVLRKAKLSVNSLSQTGGS-LYMBAGSTWDPVTPQ 555
DB 403 SEEEKTPDNLTSQLOQPIELKSGRLVKDRAVLSPSLSDQDQALLIMEAGTS----- 455

QY 556 PPOQPPAANQLITLSNLHLISLSLLANNVTPPTNPPAQDHPAVIGSTTAGSVIISGP 615
DB 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP 485

QY 616 -----IFPEDL-DDTAYDRYDMLGSKNKLNLVLKQLGTGKPPANAPSDLTL--GNEMPKY 666
DB 486 NLSIQIKIFLNSGDNFENVELLSKEQN-NIPLLT-----SKEQSHLHLPDGNLSSH 539

QY 667 GYQGSWKLAWDPNTANGPYTLKATWTCTGYNPGERVASLVPNSLWGSITLDRSAHSAI 726
```

```

Db      540 GYQGDWTFW--KDSDEG-HSLIANWTPKNYVPHPERQSTLVANTLWNTYSMDQAVQSMI 596
Qy      727 QASVDGRSYCRGLWVSGVSNFFY-HRDLALG-QGYRYISGGYSLGANSY-FGSSMFLGLAF 783
Db      597 NTTAHGGAYLFGTWSGAVSNLFYAHDSGKPIDNWHHSLGSLYLGISTHSLDDHSPCLAA 656
Qy      784 TEVGRSKDYVVCNRHHCIGSVVLSYQALCGSYLFGDAFIRASVYGFQNHKTSY-T 842
Db      657 GQLLKGSDSFITSTE-----TTSYIATVQALATSLM-KISAQACYNESIHELKTKYRS 710
Qy      843 FAESDVRDNNCLAGEICAGLPIVITPSKLYINELRPFVQAEFSYADHESFTBEGDQAR 902
Db      711 FSKEGFGSHWSVAVSGEVCAPIVNSGSLF-SFSIFSKLQSGFTQDGFESSGGEIR 769
Qy      903 APKSGHLLNLVSPGVKDFRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETWTD 962
Db      770 SFSASSFRNISLPIGITPEKKSQKTRNTYVFLGAVIQDLKRDVKSPPVTLKNAVSWDAL 829
Qy      963 AFHLARHGVVVRGSMYASITSNIEVYHGRIYRDASRGYGLSAGSRVRF 1012
Db      830 MANLDSRAYMFLTNQORAL-HRLQTLNVSYYLQGSHSYSLDLGTTVRF 878

RESULT 33
Q4VWR8 CHLTR
ID Q4VWR8 CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=HAR-13;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299427; AAQ74443.1; -; Genomic DNA.
SQ SEQUENCE 878 AA; 95538 MW; 3F3F4BFF118AEECB CRC64;

Query Match 14.8%; Score 779.5; DB 2; Length 878;
Best Local Similarity 26.6%; Pred. No. 1.9e-37;
Matches 269; Conservative 148; Mismatches 385; Indels 209; Gaps 35;

Qy      53 DPSTGTVFSAGELTKNLDNSIAALPLSCFGNLL-----GSTVLGRGHSLTPE 101
Db      26 DPLGETA-----LTKPNHVCTTFEDCTWESLFPALCAHASQDDPLVILGNSYCWFS 80
Qy      102 NIRTSTNGAALNSAADLFTIEGPKELSFNCNSLLAVLPAATNKGSTPTTTSTPSN 161
Db      81 KLHITDPEALFKEGD--LSIQNFRFLSFTDCSS-----KESSPS- 119
Qy      162 GTIYSKT-DLLLLNNEKFSFYSLNLSVGDGDAIDAKSLTVQGISKLVCVQENTAQADGAC 220
Db      120 -IIHQKQSLRNNGSMSPCRNHAHGSGGSAISADAFSLQHNLYLPTAPEENSSKNGGAI 178
Qy      221 QVTVFSMANEAPAFVANVAVRGGGIAAVQDQOQGVSSSTSTEDPVVFSFRNTAVEP 280
Db      179 QAQT-FSLSRNVSPISFARNADLNGGAIC----- 208
Qy      281 DGNVARVGGIYVSYGNVAPLNNKTLPLNNVASPVYIAAKQPTSQASNTSNNYDGGAI 340
Db      209 -----SNLCSGNV-----NPLFP-----TGNSATNGGAI 233
Qy      341 FCKNGAAGSNNGSVSGVDFGEGVFPSSNVAAGKGGAIYAKKLSVANGCPVQFLRNLIAND 400
Db      234 CC1--SDLTNBEKGLSLACQETLFAFSNAKESAKGGAIVAGHVLRYNGVPVFINNSAKI 291
Qy      401 GGAYILSGSGELSLADYGDIIIPDQNLKRTAKENADVNGVTVSSQAISMSGGKITTLR 460

```

```

Db      292 GGATATGGSLSLAGEGVLFPONNSQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343
Qy      461 AKAGHQILFNDPI-EMANGNNQPAQSSKJLKINDGEGYTG-----DIVFA----- 504
Db      344 ARNG-DILFPDPIQVESSEKESPLPSSLQASVTSPTATASPLVIQTSANRSVIFSSERL 402
Qy      505 -----NGSSTUYQNWTTIEQGRIVLREKAKLSVNSLSQGG-S-LYMEAGSTWDFVTPQ 555
Db      403 SEBEKTPDNLTSQLOPIELKSGRLVKRAVLASPSLQSDPQALLIMEAGTS----- 455
Qy      556 PPQPPAANQLITLSNLHLSSLLANNVNTPTNPQAQDSHPAVIGSTTAGSVTISGP 615
Db      456 -----LKTSSDLKIALTSLIPLHSL-----DTEKSVTTHAP 485
Qy      616 -----IPFEDL-DTAYDRYDMIGSNQKINVLKQLQTKPPANAPSDLT--GNEMPKY 666
Db      486 NLSIQKIFLNSGDNFENVVLLSKEQN-NIPLLT-----SKEQSHLHLPDGNLSSHP 539
Qy      667 GYQGSWKLAMPNTANNPGPYTLKATWTKTGNPGRVAVSLVPSNLSGSLDLSAHSAL 726
Db      540 GYQGDWTFW--KDSDEG-HSLIANWTPKNYVPHPERQSTLVANTLWNTYSMDQAVQSMI 596
Qy      727 QASVDGRSYCRGLWVSGVSNFFYHRRDALGQ--GYRYISGGYSLGANSY-FGSSMFLGLA 782
Db      597 NTTAHGGAYLFGTWSGAVSNLFY-AHDSFGKPIDNWHHSLGSLYLGISTHSLDDHSPCL 655
Qy      783 FTEVGRSKDYVVCNRHHCIGSVVLSYQALCGSYLFGDAFIRASVYGFQNHKTSY- 841
Db      656 AQQLLKGSDSFITSTE-----TTSYIATVQALATSLM-KISAQACYNESIHELKTKYR 709
Qy      842 TFAESDVRDNNCLAGEICAGLPIVITPSKLYINELRPFVQAEFSYADHESFTBEGDQA 901
Db      710 SFSKEGFGSHWSVAVSGEVCAPIVNSGSLF-SFSIFSKLQSGFTQDGFESSGGEI 768
Qy      902 RAPKSGHLLNLVSPGVKDFRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETWTT 961
Db      769 RSFSASSFRNISLPIGITPEKKSQKTRNTYVFLGAVIQDLKRDVKSPPVTLKNAVSWDA 828
Qy      962 DAFHLARHGVVVRGSMYASITSNIEVYHGRIYRDASRGYGLSAGSRVRF 1012
Db      829 PMANLASRAYMFLTNQORAL-HRLQTLNVSYYLQGSHSYSLDLGATHRF 878

RESULT 34
Q4VWR8 CHLTR
ID Q4VWR8 CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=UW-36, and UW-92;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299437; AAQ74453.1; -; Genomic DNA.
DR EMBL; AY299435; AAQ74451.1; -; Genomic DNA.
SQ SEQUENCE 878 AA; 95568 MW; BOPDF9ACB8A333AB CRC64;

Query Match 14.8%; Score 777.5; DB 2; Length 878;
Best Local Similarity 26.4%; Pred. No. 2.5e-37;
Matches 266; Conservative 145; Mismatches 394; Indels 203; Gaps 33;

Qy      53 DPSTGTVFSAGELTKNLDNSIAALPLSCFGNLL-----GSTVLGRGHSLTPE 101
Db      26 DPLGETA-----LTKPNHVCTTFEDCTWESLFPALCAHASQDDPLVILGNSYCWFS 80

```

QY	102	NIRSTNGAALNSAADGLFTIEGKELSPNSCNLSLAVLPAATNKGSTPTTTSTPSN	161
DB	81	KLHITDPKEALPKKGD--LSIQNFRPLSFDTCCS-----KESSFS-	119
QY	162	GTIYSKT-DLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVQENTAQDGGAC	220
DB	120	-IHHQNGQLSRNGSMFPCRNHAEKGGALISADAPSLQHNYLTPAENSCKNGGAI	178
QY	221	QVVTFSAMANBAPAFVANVAGVGGGIAAVODGQQGVSSSTSTEDPVVPSRNTAVEP	280
DB	179	QAQT-FSLSRNVSPISFAHRADLNGCAIC-----	208
QY	281	DGNVARVGGIYSYGNVAFPLNGKTLPLNNVASPVYIAAKQTSQASNTSNYDGGAI	340
DB	209	-----SNLICSGNV-----NLPFF-----TGSATNGGAI	233
QY	341	FCNKAQAQAGSNNSGVSFPGEGVVPFSSNVAAGKGAIIYAKKLISVANCGPVQFLRNIA	400
DB	234	CCI--SDLNTSEKLSLACNQETLPASNAKEKGAIIYAKHVLRYNGPVSPFINNSAKI	291
QY	401	GGAIYLGESGELSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAISMGSGKITTLR	460
DB	292	GGAIATQSGSLISILAGEGSLFQNNRSQRTSDQGLVR-NAIYLEKDAI-----LSSLE	343
QY	461	AKAGHOILFNDPI-EMANGNQAQSKLLKINDGEGYT-----DIVPA----	504
DB	344	ARNG-DILFPDPIVORSSSKESPLPSLQASVTPTPATASPLVIQTSANRVSIFSSERL	402
QY	505	-----NGSFLYQNVITQGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDPVTQ	555
DB	403	SEBEKTPDLNLSQLOPIELKSGRLVKORAVLSAFSLQSDPQALLIMEAGTS-----	455
QY	556	PQOPPAANQLITLSNLHLSLSLLANNAVTPPTNPPAQDSHPAVIGSTTAGSVTISGP	615
DB	456	-----LKTSSDLKLATLSLPLHSL-----DTEKSVTHAP	485
QY	616	-----IPFEDL-DDTAYDRYDWLGSNQKINVLKQLGTKPPANAPSDLTGNEMPKYGY	668
DB	486	NLSIQIFLNSGDNFYENVELLSKKQ-----NIFLLTLPKESQSHVLLPDGNLSHPGF	541
QY	669	QGSWKLAWDPNTANNPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDIRSAHSAIOA	728
DB	542	QSDWTFSW--KSDSEG-HSLIANWTPKNYVPHPERQSTLLVANTLWNTYSDMQAVQSMINT	598
QY	729	SVDRSVCRLGLWGSVSNFFY-HDRDALG-QGYRIYSGGYSLGANSY-FGSSMFLAFTPE	785
DB	599	IAHGAYLFTWGSVSNLFYAHDSCKPIDNWHRSGLYLPGLISTHSLDDHSCFLAAGQ	658
QY	786	VFGRSKDYVVCRNHHACISGVYLSLTOQALCGSYLFGDAFIRASYGPGNOHMTSY-TPA	844
DB	659	LLGKSDSPITSTE-----TTSVIATVQAQLATPLM-KISAQACYNESIHLELTKYRSFS	712
QY	845	ESDVRWNNCLAGEICAGLPVITPSKLYLNELRPVQAEFSYADHESFTESGDQARAF	904
DB	713	KEGFGSHSVAVSGEVCAPIVSVNSGSLP-SFSFISFKLQGFSGTODGFESGSEIRSP	771
QY	905	KSQHLLNLSVPVGVKPCRSSTHPNKYSFMAAYICDAYRTISGTETTLTSHQETTTTDAF	964
DB	772	SASSFRNISLPMGIFTPEKSKQNTYFYFLGAYIQDLKRDVESGPPVLLKNVSWDAPMA	831
QY	965	HLARHGVVVRGSMVASITSNIEVYGHGREYRDASRGYGLSAGSRVRF	1012
DB	832	NLDSRAYMFLTNQAL-HRLQTLNLSVYLRGQSHSYSLDLGTTTYP	878
RESULT 35			
ID	Q5L6J6	CHLAB	
AC	Q5L6J6		
DT	01-FEB-2005	(Tremblrel. 29, Created)	
DT	01-FEB-2005	(Tremblrel. 29, Last sequence update)	
DT	01-FEB-2005	(Tremblrel. 29, Last annotation update)	

DE	Polymorphic outer membrane protein.		
GN	Name=pmp10G; OrderedLocusNames=CAB277;		
OS	Chlamydia abortus.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_taxid=83555;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	STRAIN=S26/3;		
RA	Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,		
RA	Livingstone M., Cerdeno-Tarraga A.-M., Harris B., Doggett J.,		
RA	Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,		
RA	Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.,		
RT	"The Chlamydia abortus genome sequence reveals an array of		
RT	variable proteins that contribute to interspecies variation."		
RL	Genome Res. 15:629-640(2005).		
DR	EMBL; CR848038; CAH63728.1; -, Genomic_DNA.		
KW	Complete proteome.		
SQ	SEQUENCE	840 AA; 90457 MW; 8FB8CDE411AE1493 CRC64;	
Query Match			
Best Local Similarity 25.0%; Pred. No. 4.1e-37;			
Matches 260; Conservative 134; Mismatches 350; Indels 297; Gaps 34;			
QY	37	YDGETLVSPYTVIGDPSGTVFVSAGSLTKLNDNSIAALPLSCFGLNLSFTVLGRGH	96
DB	32	PDGSGTGTQAKQSTQAGGTYTNLTADVIIQHVKSTOPA-NTSCFKNSTGDTITFGANH	90
QY	97	SLTFENIHTSTNGAALSNSAADGLFTIEGKELSPNSCNLSLAVLPAATNKGSTPTTT	156
DB	91	SLIFEDIVSTAQGAISTNTDGTITMSGFNVLSF-----IAAPQATT-----	133
QY	157	STPSNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVQENTAQAD	216
DB	134	-----	133
QY	217	GCACQVVTFSAMANEAFIAFVANVAGVGRGGIAAVQDQGVSSSTSTEDPVVPSRNT	276
DB	134	-----	133
QY	277	AVEFGNVARVGGIYSYGNVAFPLNGKTLPLNNVASPVYIAAKQTSQASNTSNYGD	336
DB	134	-----GNA-----IYGIASITIKENNRV-----DTHSTAA	162
QY	337	GGAIFC-KNGAAGSNNSGVSFDCGEGVVPFSSNVAAGKGAIIYAKKLISVANCGPVQFLR	395
DB	163	GGAIHCLKTGATA-----TTLLEKNASWIFRNSSATGGAIHTDLVLTAGGYTLFEN	217
QY	396	NIAND-GGAIYLGESGELSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAISMGSGG	454
DB	218	NHATQGGAIISAGSGELSADGSGIIFRGNTYTDAGNRV-----NNAIYVANG	268
QY	455	KITTLBAKAGHOILFNDPI-EMANGNQAQSKLLKINDGEG-----YTGDIVFANGS---	507
DB	269	KPTKLEAKAQSLFYDPLIVEGAADN-----LEINKANGATTYTGSIIFSGRHTH	320
QY	508	-----STLYQNVITQGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDPVTPOPP	557
DB	321	SPHKGMKHSKFTQPLTLAAGSLILEKGAHLEAKSLUTQAGRIILDQTS-----	371
QY	558	QQPAAANQLITLSNLHLSLSLLANNAVTPPTNPPAQDSHPAVIGSTTAGSVTISGPI-	616
DB	372	---IKVQENVDIKDLWLSLDFEVEPTA-----TH--IASSGDHNVITIGPLG	414
QY	617	FPEDLDDTAYDRYDWLGSNQKINVLKQLGTKPPANAPSDLTGN--EMPK-----	665
DB	415	IPAD-QETFPYHHA-LAHNVDBELLQL-----ADKDIKISLVDVPESVRENKDAH	463
QY	666	YGVQGSWKLAWD--PNTANN-----GPYTLKATWTKTGNP--GPERVAS-LVPNSLWGS	715
DB	464	RGYQGSWTIDWKTVPGSTSGGVTLGTATVHWRPTGYIFPGGAQEIITPLVPNTLWGN	523
QY	716	ILDIRSAHSAIQASVDGRSYCRGLWGSVSNFPFHDRLALGQGYRYISGGYSLGANSY-F	774



```

Db 524 FSDINRLERTIESLATNSLCSGFWATGIKFLYTSQGEKFPVQHNNSGVGAIGINKHTL 583
Qy 775 GSSMFLAFTVFRGSKDYVYVCRNHHACISGVY---LSTQALCGSYLFGDA----- 824
Db 584 SENVFAAFSQLFGKDRDQAQOEHQTLGSLYAHVGTIPML--REFCGDSKQCVPEL 641
Qy 825 -----FIRASVQFGNHKMTSYTFABESDVR-----WDNNCLAGEAGLPIVITPSK 872
Db 642 QASPCIPVIFNAQLSYSHS--NNLTIAHEDQTKTTGNWNSYVATELGSTFVYTLKCP 699
Qy 873 LYLNELRFVQAEPSYADHESFTREGDQARAFKSGHLLNLSPVGVKFRDCSSTHPNKYS 932
Db 700 SILKNSPFKLGQVYSQRATFEGLRCLFSSYLANLALPVGIKIGICPKRLAYD 759
Qy 933 FMAAYICDAYRTIGTETT--LLSHQETWTTTDAFLARHGVVVRGSMVASLTSNIEVYGHG 991
Db 760 LSAMYVHDVFRINPETMTLFLIGRLAPWTTTATHLDNKALVQSGRFAVRPNIEVPASG 819
Qy 992 RYERDASRGYGLSAGSVRF 1012
Db 820 NGELRSSSHSYNDFGAKIHF 840

```

## RESULT 36

```

Q84FV4 CHLTR
ID Q84FV4 CHLTR PRELIMINARY; PRT; 846 AA.
AC Q84FV4;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Polymorphic membrane protein I (fragment).
GN Name=pmpl;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22483673; PubMed=12595433;
RX DOI=10.1128/IAI.71.3.1200-1208.2003;
RA Stothard D.R., Toth G.A., Battiger B.B.;
RT "Polymorphic membrane protein H has evolved in parallel with the three
RT disease-causing groups of Chlamydia trachomatis."
RL Infect. Immun. 71:1200-1208 (2003).
DR EMBL; AY184171; AAC029988.1; -; Genomic DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto transportbeta.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter_1.
DR Pfam; PF07548; ChlamPMP M; 1.
DR Pfam; PF02415; Chlam PMP; 6.
DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
FT NON_TER 1
FT TER 846
FT NON_TER 846
FT SEQUENCE 846 AA; 92007 MW; B0E2F7ED26A75ED CRC64;

```

```

Query Match 14.6%; Score 771.5; DB 2; Length 846;
Best local similarity 26.4%; Pred. No. 5.5e-37;
Matches 264; Conservative 145; Mismatches 388; Indels 203; Gaps 33;
Qy 53 DPGTTFVPSAGELLKNDLNSIALPLSCFQNL-----GSTVLGRGHSITYPE 101
Db 2 DPLGETA-----LTKPNHVCTTFEDCTMESLPALCAHQSDDPLVYLGNSYCFVS 56
Qy 102 NIRTSTGAALNSAAGDLFTIEGKELSPNCNLLAVLPAATTNKGSTPTTTSTPSN 161
Db 57 KLHTDPKEALFKKGD--LSIQNFRFLSFTDCSS-----KESSFS- 95
Qy 162 GTIYSKT--DLLLANNKFSFYSNLSVGGDAIDAKSLTVQGISKLCVFPQNTAQDGGAC 220
Db 96 -IIHQKQLSLRNNGSMFCRNEAGSGGGAISADAFSLQHLYLFTAPEENSKNGGAI 154

```

```

Qy 221 QVVTSPSAMANEAPIAFVANVAGVGGIIAAVQDQGGQSSSTSTEDPVVFSRNTAVEF 280
Db 155 QAQT-FSLSRNVSPISFARNADLNGGAICC----- 184
Qy 281 DGNVARVGGIGYSYGNVAFNLNGKTLPLNNVASPVYIAAKQPTSGQASNTSNNTGDDGAI 340
Db 185 -----SNLICSGNV-----NPLFF-----TGNBATNGAI 209
Qy 341 FCNKAQAQAGNNGSVSFGDGVVFFSSNVAAGKGAIAKLSVANCGVPQVFLURNIAND 400
Db 210 CCI--SDLNTSEKGLSLACNQEITLPASNAKEKGAIAKHMVLRVNGPVSPFINNSAKI 267
Qy 401 GGAIIYGESELSADYDGIIFDGNLKRTAKENADVNGVTSSQIASMGSGKIIITLR 460
Db 268 GGAIAIQSGSLSILAGEBSVLFPNNRSQRTSDQGLVR-NAIYLEKDAI-----LSSLE 319
Qy 461 AKAGHQILFDNPT-EMANGNNOQAQSSKLLKINDGEGYTG-----DIVFA--- 504
Db 320 ARNG-DILFPDPIVOESSKESPLPSSLOASVSTPTATASPLVIQTSANRSVIFSSERL 378
Qy 505 -----NGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
Db 379 SEBEKTPDNLTSOLQQPIELKSGRLVKDRAVLSPSLSDQPQALLIMEAGTS----- 431
Qy 556 PPOQPPAANQIITLSNLHLSLSLLANNVNTPTNPQAQDSHPAVIGSTAGSVTISGP 615
Db 432 -----LKTSSDLKATLSIPLHSL-----DTEKSVTHIAP 461
Qy 616 -----IFFEDL-DDTAYDRYDWMISNOKINVLKQLGKTKPPANAPSDLTGNEMPKYGY 668
Db 462 NLSIQKIFLNSGDNFYNVELLSKEQN-----NIPLLTLPKQSHLHLPDGNLSSHPGY 517
Qy 669 QGSWKLAWDNPNTANGPYTLKATWTKTGYNPGPERVASLVPNSLWGLSILDIRSAHSAIOA 728
Db 518 QGDWTFSM--KDSDEG-HSLANWTPKVYVHPERQSTLVANTLWNTYSDMQAVQSMINT 574
Qy 729 SVDGRSYCRGLWVGVSNNPFY-HDRDALG-QGYRTISGGYSLGAN8Y-FCSSMPLGAPTE 785
Db 575 TAHGGAYLFGTWGSAVSNLFYAHDSGKPIDNWHHRSGLYLFGISTHSLDDHDFCLAAQ 634
Qy 786 VFGRSKDYVYVCRNHHACISGVYLSLTOALCGSYLFGDAFIRASYPGNQHKMTSY-TFA 844
Db 635 LLGKSSDPTSTTE-----TTSYIATVQAOLATSLM-KISAQACYNESIHELKTKYRSFS 688
Qy 845 EESDVRMDNCLAGEIGAGLPVITPSKLYNELRPFVQAEFSYADHESFTREGDQARAF 904
Db 689 KEGFGSMHSAVSGEVCALIPVNSGSLF-SSFSIFSKLQGFSGTQDGPESSGEIRSF 747
Qy 905 KSGHLLNLSPVGVKPRDCSSTHPNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAF 964
Db 748 SASSFRNISLPIGITTFEKKSQKTRNYYFYGAVIQDLKRDVSGPVVLLKNVSWDAPNA 807
Qy 965 HLAHGVVVRGSMVASYLTSNIEVYHGRYERDASRGYGL 1004
Db 808 NLASRAYMFLTNQRAL-HRLQTLNLNVSVLRQGSYSYL 846

```

## RESULT 37

```

PMPI CHLTR
ID PMPI CHLTR STANDARD; PRT; 878 AA.
AC O84882;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable outer membrane protein pmpl precursor (Polymorphic membrane
DE protein 1).
GN Name=pmpl; OrderedLocusNames=CT874;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=D/UW-3/Cx;

```

EX MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;  
RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.B., Zhao Q., Koonin E.V.,  
RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans:  
Chlamydia trachomatis";  
RL Science 282:754-759(1998).  
CC -!- SURCELLULAR LOCATION: Cell wall surface (elementary bodies)  
CC (potential).  
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.  
CC  
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC -----  
CC EMBL: AB001361; AAC68472.1; -; Genomic DNA.  
CC PIR: B71460; B71460.  
CC InterPro: IPR005546; Auto transpbeta.  
CC InterPro: IPR011427; ChlamPMP\_M.  
CC InterPro: IPR003368; Chlamydia\_PMP.  
CC Pfam: PF03797; Autotransporter; 1.  
CC Pfam: PF02415; ChlamPMP; 6.  
CC Pfam: PF07548; ChlamPMP\_M; 1.  
CC TIGRfam: TIGR01376; POMP\_repeat; 6.  
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.  
FT SIGNAL 1 24 Potential.  
FT CHAIN 25 878 Probable outer membrane protein pmp1.  
FT SEQUENCE 878 AA; 95593 MW; DF1F1A31707EE48B CRC64;  
SQ  
  
Query Match 14.6%; Score 771.5; DB 1; Length 878;  
Best Local Similarity 26.5%; Pred. No. 5.8e-37;  
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;  
  
QY 53 DPSTGVFSGELTAKLNDNSIAALPLSCFGLN-----GSFTVLGRGHSLTPE 101  
DB 26 DPLGETA-----LTKNPNHVCTTFEDCTMESLFPALCNHASQDDPLVLGNSYCFVS 80  
  
QY 102 NIRTSTNGAALNSAADGLFTIEGPKELSFNCNLSLLAVLPAATTNKGSTPTTTSTPSN 161  
DB 81 KLHITDPKEALPKBKGD--LSIQNPRFLSPDCCS-----KESSPS- 119  
  
QY 162 GTIYSKT-DLLLNNEKFPYSNLVSGDGAIDAKSLTWQISKLCVFOENTAAQAGGAC 220  
DB 120 -IIHQKNGQLSLRNGSMSPCRNHAEGSGAIGADAFSLQHNYLFTAFENSKGGGAI 178  
  
QY 221 QVVTSSAMANEAPIAFVANVAGVGGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVPE 280  
DB 179 QAQT-FSLSRNVSPISFARNRADLNGAICC-----NPLPP----- 208  
  
QY 281 DGNVAVGGGIYGYGNVAFNLNGKTLFLNNVAPSVVIAAKQPTSGQASNTSNNGYDGGAI 340  
DB 209 -----SNLIGENV-----NPLPP-----TGN SATNGGAI 233  
  
QY 341 FCKNGAAGSNNGSVSFDGEGVFPSSNVAAGKGAIAKLVANCGVQFLRIAND 400  
DB 234 CCI--SDLNTSEKSLSLACNQETLFAASNAKKGGAIAKHMVLRNGVPVFINNSAKI 291  
  
QY 401 GGAIIYGESGELSLSADYDIIIDGNLKRKTAKENADVNGVTVSSQASISMGSGGKITTLR 460  
DB 292 GGAITAQSGGSLSLAGEGSLVFQNNNQSTQDGLVR-NAIYLEKDAI-----LSSLE 343  
  
QY 461 AKAGHQILFNDPT-EMANGNNQPAQSSKLKINDGEGYTG-----DIVFA----- 504  
DB 344 ARNG-DILFDPPIVOESSEKESPLPSSLQASVTSPTATAPLVIQTSANRVSFSSERL 402  
  
QY 505 -----NGSSTLYQNVTTIEQGRIVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFTPQ 555  
DB 403 SEBEKTPDNLTSQLOQPIELKSGRLVKDRVLSAPLSQDPOALLIMEAGTS----- 455  
  
QY 556 PPOQPPAANQLITLSNLHLSLSLLANNVNTNPTNPQADSHPAVIGSTAGSVTISGP 615

DB 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHIAP 485  
QY 616 -----IPEFDL-DDTAYDRYDMLGSKQKINVLKLGTKPPANAPSDLTL--GNEMPKY 666  
DB 486 NLSQKIFLSNGSDENFYENVLLSKQN-NIPLTL-----SKEQSHLHLPDGNLSSH 539  
QY 667 GYQGSWKLAWDPNTANNNGPYTLKATWTKTGYNPGRVASLVPNSLWGSILDIRSAHSAI 726  
DB 540 GYQGDWTFW--KDSDEG-HSLIANWTPTKNTVPHPERQSTLVANTLWNTYSDMQAVQSMI 596  
QY 727 QASVDGRYCGELWVGVSNEFY-HDRDALG-QGVRYISGGVSLGANYS-FGSSMFGGLAF 783  
DB 597 NTIAHGAYLFGTWGSAVSNLFYADHSSGKPIDNWHHSLGLVLFISTHSDHDFCLAA 656  
QY 784 TEVFGSRKDYVVCNHHACIGSVVLSYQOALCGSYLFGDAFIRASYGFGNQHMKTSY-T 842  
DB 657 GQLLKKSSDSFITSTE-----TTSVIATVQAQLATPLM-KLSAQACYNESIHELKTKYS 710  
QY 843 FAESDVRWNNCLAGEIGAGLPVITPSKLYNELRPFVQABFSAADHESPTERGDOAR 902  
DB 711 FSKGFGSWHSVAVSGEVCASPIVNSGGLP-SSFSIFSKLQGFSGTQDGFESSGEIR 769  
QY 903 AFKSGHLNLNLSVPVGVKEDRCSSHTPNKYSFMAAYICDAYTISCTETLLSHQETWTD 962  
DB 770 SFSASSFRNISLPMGITTEKKSQKTRNYYPFLGAYIQDLKRDVRESGPPVLLKNVSWDAP 829  
QY 963 AFHLARHGVRGSMYASILTSNIEVYGHGRVEYRDASRGYGLSAGSRVRF 1012  
DB 830 MANLDSRAYMFLTNQRAL-HRLQTLNLSVYVLRQSHSYSLDLGTTYRF 878  
  
RESULT 38  
Q4VMS2\_CHLTR PRELIMINARY; PRT; 878 AA.  
ID Q4VMS2;  
AC Q4VMS2;  
DT 13-SEP-2005 (TREMREL. 31, Created)  
DT 13-SEP-2005 (TREMREL. 31, Last sequence update)  
DT 13-SEP-2005 (TREMREL. 31, Last annotation update)  
DE Polymorphic membrane protein I.  
GN Name=pmp1;  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=UW-12, UW-202, UW-31, UW-4, and UW-3;  
RA Gomes J.P., Borrego M., Dean D.;  
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY299433; AAQ74449.1; -; Genomic DNA.  
DR EMBL: AY299434; AAQ74450.1; -; Genomic DNA.  
DR EMBL: AY299436; AAQ74452.1; -; Genomic DNA.  
DR EMBL: AY299431; AAQ74447.1; -; Genomic DNA.  
DR EMBL: AY299445; AAQ74461.1; -; Genomic DNA.  
DR InterPro: IPR005546; Auto transpbeta.  
DR InterPro: IPR011427; ChlamPMP\_M.  
DR InterPro: IPR003368; Chlamydia\_PMP.  
DR Pfam: PF03797; Autotransporter; 1.  
DR Pfam: PF02415; ChlamPMP; 6.  
DR Pfam: PF07548; ChlamPMP\_M; 1.  
DR TIGRfam: TIGR01376; POMP\_repeat; 6.  
SQ SEQUENCE 878 AA; 95593 MW; DF1F1A31707EE48B CRC64;

Query Match 14.6%; Score 771.5; DB 2; Length 878;  
Best Local Similarity 26.5%; Pred. No. 5.8e-37;  
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;  
  
QY 53 DPSTGVFSGELTAKLNDNSIAALPLSCFGLN-----GSFTVLGRGHSLTPE 101  
DB 26 DPLGETA-----LTKNPNHVCTTFEDCTMESLFPALCNHASQDDPLVLGNSYCFVS 80  
  
QY 102 NIRTSTNGAALNSAADGLFTIEGPKELSFNCNLSLLAVLPAATTNKGSTPTTTSTPSN 161

```

Db      81 KLHITDPKEALFKEGD--LSIQNFRPLSFPTDCS-----KESSPS- 119
Qy      162 GTIYSKT-DLLLNNEKSFYNLVSGDGGAIADAKSLTVQGISKLCVFOENTAOADGAC 220
Db      120 -IIHQKQGLSLRNNGSMSCFNHAGSGGAIADAFSLQHNYLFTAFENSSKNGGAI 178
Qy      221 QVVTSPSAMANEAPAFANVAVGVRGGGIAAVQDQCGQVSSSTSTEDPVVFSRNTAVEF 280
Db      179 QAQT-FSLSRNVSPISFARNRADLNGGAI--LSIQNFRPLSFPTDCS-----KESSPS- 119
Qy      281 DGNVAVGGIYSGYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNTSNNYDGGAI 340
Db      209 -----SNLICSGNV-----NPLFF-----TGN SATNGAI 233
Qy      341 FCNKAQAGSNNGSVSGDGGVFPSSNVAGKGGAIYAKKLSVANGCPVQFQRLNAND 400
Db      234 CCI--SDLNTSEKGLSLACNQETLFAASNAKEKGGAIYAKHMLVRYNGPVSFNNNAKI 291
Qy      401 GGAIYLGESGELSLADYDGIIFDGNLRTAKENADVNGVTVSSQAISSMGSGGKITTLR 460
Db      292 GGAIYQSGGSLILAGSGVLQNNQSRTSDQGLVR-NAIYLEKDAI-----LSLE 343
Qy      461 AKAGHQILFNDPI-EMANGNPOAQSSKLLKINDGEGYT-----DIVFA----- 504
Db      344 ARNG-DILFPDPIVQESSKESPLPSLSQASVTSPTATASPLVIQTSANRSVIFSSERL 402
Qy      505 -----NGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQ 555
Db      403 SEBEKTPDNLTSQLOQPIELKSGRLVKDRAVLSAPLSQDPQALLIMEAGTS----- 455
Qy      556 PPQPPAANQLITLSNLHLSLLANNVTPNPPAQDSHPAVIGTSITAGSVTISGP 615
Db      456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP 485
Qy      616 -----IPFEDL-DDTAYDRYDMGLSNQKINVLKQLGKTPPANAPSDLTL--GNEMPKY 666
Db      486 NLSIQKIFLNSGDNENFENVLLSKEQN-NIPLLT-----SKEQSHLHLPDGNLSHP 539
Qy      667 GYQGSWKLAWDPNTANNPPTLKATWTKTGNYPGERVASLVPNSLWGSILDIRSAHAI 726
Db      540 GYQGDWTFPSW--KDSDEG-HSLIANWTPKNYVPHPERQSTLVANTLNTYSDMQVQSMI 596
Qy      727 QASVDGRSYCRGLWVGVSNNFFY-HDRDALG-QGYRYISGGYSLGANSY-FGSSMFGIAP 783
Db      597 NTIAHGGAFLFTWGSVAVSNLIFYADSSGKPIDNWHRSLOYLFGISTHSDHSCFLAA 656
Qy      784 TEVGRSKDYVVCNRRHHACIGSVYLSLTOQALCGSYLFGDAFIRASYGFGQHMKTSY-T 842
Db      657 QOLLGKSSDSFITSTE-----TTSYIATVQALATPLM-KISAQACYNESITHELTKYRS 710
Qy      843 PABSDVWMDNCLAGEIGAGLPVITPSKLYLNLPRFPVQAFSYADHESFTBEGDOAR 902
Db      711 FSKEGFGSWHSVAVSGEVCASIPVNSGSLF-SFSFISFKLQFGSGTQDGFESSGGEIR 769
Qy      903 AFKSHLLNLVSPVGVKEDRCSSHPNKYSFMAAYICDAYTIGTETITLLSHOETWTD 962
Db      770 SFSASSFNISLPMGITEKKSQKTRNYTYFLGAYIQDLKRDVSGPVVLLKNAVSWDAP 829
Qy      963 APHLARHGVVVRGMSVYASLTNIEVYGHRYEYRDASRGYGLSAGSRVF 1012
Db      830 MANLDSRAYMFLTNQORAL-HRLQTLNLVSVYLRGQSHSYSLDGLTGYTRF 878

```

## RESULT 39

```

Q4VWR1.CHLTR
ID Q4VWR1.CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR1.
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Polymorphic membrane protein 1.
GN Name=pmp1;
OS Chlamydia trachomatis.

```

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI\_TaxID=813;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=UW-57;

RA Gomes J.P., Borrego M., Dean D.;

RT \*Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.\*;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY299444; AAQ74460.1; -, Genomic DNA.

SQ SEQUENCE 878 AA; 95534 MW; DF1FIA333B13501B CRC64;

Query Match 14.6%; Score 770.5; DB 2; Length 878;

Best Local Similarity 26.5%; Pred. No. 6.6e-37;

Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

Qy 53 DPSGTTVFSAGELTLKNLDSIALPLSCFQNL-----GSFTVLGRGSLTTFE 101

Db 26 DPLGETA-----LITKPNHVCTFFEDCTMESLPALCAHASODDPLVVLGNSYCMFVS 80

Qy 102 NIRTSTGAALSNSAADGLFTIEGKELSPNSCNLSLAVLPAAATNKGSTPTTTSTPSN 161

Db 81 KLHITDPKEALFKEGD--LSIQNFRPLSFPTDCS-----KESSPS- 119

Qy 162 GTIYSKT-DLLLNNEKSFYNLVSGDGGAIADAKSLTVQGISKLCVFOENTAOADGAC 220

Db 120 -IIHQKQGLSLRNNGSMSCFNHAGSGGAIADAFSLQHNYLFTAFENSSKNGGAI 178

Qy 221 QVVTSPSAMANEAPAFANVAVGVRGGGIAAVQDQCGQVSSSTSTEDPVVFSRNTAVEF 280

Db 179 QAQT-FSLSRNVSPISFARNRADLNGGAI--LSIQNFRPLSFPTDCS-----KESSPS- 119

Qy 281 DGNVAVGGIYSGYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNTSNNYDGGAI 340

Db 209 -----SNLICSGNV-----NPLFF-----TGN SATNGAI 233

Qy 341 FCNKAQAGSNNGSVSGDGGVFPSSNVAGKGGAIYAKKLSVANGCPVQFQRLNAND 400

Db 234 CCI--SDLNTSEKGLSLACNQETLFAASNAKEKGGAIYAKHMLVRYNGPVSFNNNAKI 291

Qy 401 GGAIYLGESGELSLADYDGIIFDGNLRTAKENADVNGVTVSSQAISSMGSGGKITTLR 460

Db 292 GGAIYQSGGSLILAGSGVLQNNQSRTSDQGLVR-NAIYLEKDAI-----LSLE 343

Qy 461 AKAGHQILFNDPI-EMANGNPOAQSSKLLKINDGEGYT-----DIVFA----- 504

Db 344 ARNG-DILFPDPIVQESSKESPLPSLSQASVTSPTATASPLVIQTSANRSVIFSSERL 402

Qy 505 -----NGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQ 555

Db 403 SEBEKTPDNLTSQLOQPIELKSGRLVKDRAVLSAPLSQDPQALLIMEAGTS----- 455

Qy 556 PPQPPAANQLITLSNLHLSLLANNVTPNPPAQDSHPAVIGTSITAGSVTISGP 615

Db 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP 485

Qy 616 -----IPFEDL-DDTAYDRYDMGLSNQKINVLKQLGKTPPANAPSDLTL--GNEMPKY 666

Db 486 NLSIQKIFLNSGDNENFENVLLSKEQN-NIPLLT-----SKEQSHLHLPDGNLSHP 539

Qy 667 GYQGSWKLAWDPNTANNPPTLKATWTKTGNYPGERVASLVPNSLWGSILDIRSAHAI 726

Db 540 GYQGDWTFPSW--KDSDEG-HSLIANWTPKNYVPHPERQSTLVANTLNTYSDMQVQSMI 596

Qy 727 QASVDGRSYCRGLWVGVSNNFFY-HDRDALG-QGYRYISGGYSLGANSY-FGSSMFGIAP 783

Db 597 NTIAHGGAFLFTWGSVAVSNLIFYADSSGKPIDNWHRSLOYLFGISTHSDHSCFLAA 656

Qy 784 TEVGRSKDYVVCNRRHHACIGSVYLSLTOQALCGSYLFGDAFIRASYGFGQHMKTSY-T 842

Db 657 QOLLGKSSDSFITSTE-----TTSYIATVQALATPLM-KISAQACYNESITHELTKYRS 710

Qy 843 PABSDVWMDNCLAGEIGAGLPVITPSKLYLNLPRFPVQAFSYADHESFTBEGDOAR 902

Db 711 FSKEGFGSWHSVAVSGEVCASIPVNSGSLF-SFSFISFKLQFGSGTQDGFESSGGEIR 769

Db	711	PSKEGFGSWHSAVGEVCASIPVYNSGLP--SSFSIFSKLQFGSQTDGPFSSGSEIR	769
Qy	903	AFKSGHLLNLSVPVGVKFDRCSTTHPNKYSFMAAICDAYRTISGTETLLSHQETWTD	962
Db	770	SFSASSFRNISLPMGITPEKSKQTRNYYIFLGAYIQDLKRDVESGVPVLLKNVSWGAP	829
Qy	963	AFHLARHGTVVGRGVSAYISLTSNIEVVGHRVEYRDASRGYGLSAGSRVRF	1012
Db	830	MANLDSRAYMFLTNQRAL--HRLQTLNLSVYVLRGQSHSYSLDLGTTTFP	878
RESULT 40			
Q84FU2_CHLTR			
AC	Q84FU2;	PRELIMINARY; PRT; 846 AA.	
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Polymorphic membrane protein I (Fragment).		
GN	Name=pmpI;		
OS	Chlamydia trachomatis.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=813;		
RV	[1]		
RN	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22483673; PubMed=12595433;		
RX	DOI=10.1128/JAI.71.3.1200-1208.2003;		
RA	Stothard D.R., Toth G.A., Basteiger B.E.;		
RT	"Polymorphic membrane protein H has evolved in parallel with the three		
RT	disease-causing groups of Chlamydia trachomatis.";		
RL	Infect. Immun. 71:1200-1208(2003).		
DR	EMBL; AY184173; AAC029990.1; -; Genomic DNA.		
DR	GO; GO:0019867; C:outer membrane; IEA.		
DR	InterPro; IPR005546; Auto transportbeta.		
DR	InterPro; IPR011427; ChlamPMP M.		
DR	InterPro; IPR003368; Chlamydia_PMP.		
DR	Pfam; PF03797; Autotransporter; 1.		
DR	Pfam; PF07548; ChlamPMP M; 1.		
DR	Pfam; PF02415; Chlam PMP; 6.		
DR	TIGRfam; TIGR01376; POMP_repeat; 6.		
FT	NON_TER 1 1		
FT	NON_TER 846 846		
SEQ	SEQUENCE 846 AA; 92023 MW; 0A50AF8E507E8FP6 CRC64;		
Query Match 14.6%; Score 769.5; DB 2; Length 846;			
Best Local Similarity 26.4%; Pred. No. 7.2e-37;			
Matches 264; Conservative 145; Mismatches 388; Indels 203; Gaps 33;			
Qy	53	DPGTTVFSAGELTKNLDNSIALPLSCFGNLL-----GSFTVLGRGHSITPE	101
Db	2	DPLGETA-----LITKPNHVCTTFEDCTMESLPALCAHSAQDDPLYLGNISYCFVS	56
Qy	102	NIRTSTNGAALNSAAGLPTIEGFKELSPNCNLSLAVLPAAATNKGSTPTTTSPTSN	161
Db	57	KLHITDPKEALFKEKGD--LSIQNFRFLSFDCSS-----KESFPS-95	
Qy	162	GTIYSKT-DLILLNNEKFSYSLNLSVDGGDAIDAKSLTVQGISKLVPQENTAOADGAC	220
Db	96	-IHKNGQLSLRNGSNMFCENHAEGSGGAI SADAFSLQHNYLFTAFENSSKNGGAI	154
Qy	221	QVWTSFSAANEAFIAFVANVAVRGGGIAAQVQGGQGVSSSTEDPVVPSFRTAVEP	280
Db	155	QAQT-FSLSRNVSPISFSRNRADLNGGAICC-----	184
Qy	281	DGNVARVGGGIYSYGNVAFINNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGDGAI	340
Db	185	-----SNLICSGNV-----NPLFF-----TGNATNGGAI	209
Qy	341	PKNGAAGSNNSGVSPDGGVVPFSSNVAAGKGAIYAKKLSVANGCPVQFLNTIAND	400
Db	210	CCI--SDLNTSEKSLSLACNQETLFPASNAKEKGAIYAKHMLRYNGPVSPFINNSAKI	267
Qy	401	GGAIYLGESGSLSLADYGDIIFDGNLKR TAKENAADVNGVTVSSQAISMGSGGKIITLR	460

Db	268	GGATAIQSGGSLSLAGEGSLFQWNSQRTSDQGLVR--NAILYLEKDAI-----LSSLR	319
Qy	461	AKAGHQILLFNDPI--EMANGNNQPAQSSKLLKINDGEGYTG-----DIVFA----	504
Db	320	ARNG--DILFFPDI VOESSKESPLPSSLOASVTSPTPATASPLVITQTSANRSVIFSSERL	378
Qy	505	-----NGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGGG--LYMEAGSTWDFVTPQ	555
Db	379	SEEEKTPDNLTSQLQPIELKSGRLVKDRAVLSAPLSQDPQALLIMEAGTS-----	431
Qy	556	PPQPPAANQLITLNLHLSLSSLLANNVTPNPPPAQDSHPAVIGSTTAGSVTISGP	615
Db	432	-----LKTSSDLKATLSLPLHSL-----DTEKSVTIHAP	461
Qy	616	-----IPEDL-DDTAYDRYDMLGSNOKINVLKQLGTGKPPANAPSDLTLGNEMPKYGY	668
Db	462	NLSIQIKFILNSGGDENFENVELLSKEQN-----NIPLLTLPKESQHLHLPGDNLSSHFGY	517
Qy	669	QGSWKLAWDPNTANNPGPYTLKATWTKTYNGPGERVASLVPNSLWGSILDIRSAHAIQA	728
Db	518	QGDWTFW--XDSDEG--HSLIANWTPKNVPHPERQSTLVANTLWNTYSDMQAVQSMINT	574
Qy	729	SVDGRSYCRGLWVGVSNNPFY--HRRDALG--QGYRYISGYSILGANSY--PGSSMFGIAPTE	785
Db	575	TAHGGAYLFGTWGSAVSNLFYAHDSGSKFPIDNWHRSLSGLYLFGISLTHSLDDHSPCLAAQ	634
Qy	786	VFRSKDYVVCNSHHACIGSVYVLSTQOALCGSYLFGDAFIRASYGFGNQHMKTSY--TFA	844
Db	635	LLGKSDSDSPITSTE-----TTSYIATVQALATSLM--KLSAOCAYNESIHELKTKYRSFS	688
Qy	845	BESDVRDNNCLAGBITGAGLPITVITPISKLYLNLRLPFFVOAEFSYADHESFTTEGDOARAF	904
Db	689	KEGFGSWHSAVSGEVCALIPVNSGSLF--SSFSIFSKLQFGSQTDGPFSSGBIRSF	747
Qy	905	KSGHLLNLSVPVGVKFDRCSTHPNKYSFMAAICDAYRTISGTETLLSHQETWTDAP	964
Db	748	SASSFRNISLPIGITPEKSKQTRNYYIFLGAYIQDLKRDVSGPVVLLKNVSWDAPWA	807
Qy	965	HLARHGTVVGRGVSAYISLTSNIEVVGHRVEYRDASRGYGL	1004
Db	808	NLASRAYMFLTNQRAL--HRLQTLNLSVYVLRGQSHSYSL	846

Search completed: May 13, 2006, 12:19:49  
Job time : 263 secs

**THIS PAGE BLANK (USPTO)**